SCIENCE LIBRARY

JUN 3 1935

Vol. XVIII, No. 3 *PSYCHOLOGICAL REVIEW PUBLICATIONS* June, 1935

18

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA

WALTER S. HUNTER, CLARK UNIVERSITY (*Psychol. Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Psychol. Rev.*)
JOSEPH PETERSON, GEORGE PEABODY COLLEGE (*Psychol. Monog.*)
JOHN A. McGEOCH, UNIVERSITY OF MISSOURI (*Psychol. Bull.*)

CONTENTS

An Action Potential Study of Neuromuscular Relations: S. R. HATHAWAY .. 285

The Conditions of Verbal Configuration: H. L. HOLLINGWORTH 299

An Attempt to Condition Hand Withdrawal Responses in Human Subjects: H. CASON 307

Reading and Legibility of Chinese Characters: IV. An Analysis of Judgments of Positions of Chinese Characters by American Subjects: S. K. CHOU .. 318

The Biochemical Variability of the Individual in Relation to Personality and Intelligence: H. GOLDSTEIN 348

The Quantitative Measurement of Laterality: C. VAN RIPER 372

A Comparison of Five Methods of Scoring the Galvanic Skin Response: W. A. HUNT AND E. B. HUNT 383

Discussion:

A Further Note on Learning without Awareness of What is Being Learned: E. L. THORNDIKE AND R. T. ROCK, JR. 388

A Rejoinder: F. W. IRWIN 389

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

PUBLICATIONS

OF THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITED BY

S. W. FERNBERGER, University of Pennsylvania (*J. Exper. Psych.*)
WALTER S. HUNTER, Clark University (*Index and Abstracts*)
HENRY T. MOORE, Skidmore College (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, Princeton University (*Review*)
EDWARD S. ROBINSON, Yale University (*Bulletin*)
JOSEPH PETERSON, George Peabody College (*Monographs*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the Bulletin consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bi-monthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 900 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The Index is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The Monographs appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, April, July, October, January, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: $5.50 (Foreign, $5.75). **Index:** $4.00 per volume.
Journal: $7.00 (Foreign, $7.25). **Monographs:** $6.00 per volume (Foreign, $6.30).
Bulletin: $6.00 (Foreign, $6.25). **Abstracts:** $6.00 (Foreign, $6.25).
Abnormal and Social: $5.00 (Foreign, $5.25). Single copies $1.50.
Current numbers: **Journal,** $1.25; **Review,** $1.00; **Abstracts,** 75c; **Bulletin,** 60c.

COMBINATION RATES

Review and Bulletin: $10.00 (Foreign, $10.50).
Review and J. Exp.: $11.00 (Foreign, $11.50).
Bulletin and J. Exp.: $12.00 (Foreign, $12.50).
Review, Bulletin, and J. Exp.: $16.00 (Foreign, $16.75).
Review, Bulletin, J. Exp., and Index: $19.00 (Foreign, $19.75).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, NEW JERSEY

Journal of
Experimental Psychology

VOL. XVIII, No. 3 JUNE, 1935

AN ACTION POTENTIAL STUDY OF NEUROMUSCULAR RELATIONS

BY STARKE R. HATHAWAY[1]

University of Minnesota, Minneapolis, Minnesota

The following experiments have been performed in the hope that they may shed light upon the functioning of the intact organism.

There is a tendency to stretch the reflex arc concept to the point where it may be offered as a functional unit which is the unit not only of overt behavior but also of implicit behavior. if what was originally called central orientational behavior Is to be explained under this concept, then it must be assumed, as it has been, that the muscles form a part of this implicit reflex circuit. The effector side of such a system is often assumed to include faint contractions of the specific muscles involved; the specificity of these faint contractions to the implicit activity is further assumed—otherwise it would be necessary to postulate that a contraction in a comparatively minor muscle stands for, or assumes, a major part in the reaction of a total system employing mainly the specific muscle of the reaction.

The presence of these faint contractions, particularly in the throat muscles, has been sought for by many investigators,

[1] Condensed from a thesis presented to the Graduate School of the University of Minnesota in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

but the search has been only partially successful It has been. demonstrated that accompanying certain central states, particularly of naïve subjects, muscular movements may be observed or recorded by sensitive apparatus. In such studies, the apparatus used has not always been of sufficient sensitivity to demonstrate as small a unit as a single neuron operating with its attendant discharges in a number of muscle cells. Since the investigators realized that their apparatus was not sensitive enough to show minimum contractions, the absence of records in certain cases may have indicated the lack of sufficient sensitivity to record a contraction which might still be assumed as present.

It has been demonstrated beyond doubt that imagination or other conscious implicit motor activities or sets may be accompanied by specific muscular contractions of a minute order. This has been most recently demonstrated in the work of Jacobson (1930*ab*, 1931) and Freeman (1931). Jacobson concludes definitely that 'motor imagination' is accompanied by action potentials, but his data show only that such action potentials may accompany it since not all his records showed potentials. He further states no evidence as to whether the action potentials mirror the movement imagined, or whether they are simply present in the general region of that muscle which the subject imagines as contracting. Though he does not state the point, he seems to imply a dependence of the imagined content upon the muscular changes. Freeman, however, claims only to demonstrate a general muscular change in tone and does not treat the problem of whether this change is secondary or primary to the set preceding the act. He does find generalized changes in muscular tension developing within the warning period preceding the stimulus of a simple reaction situation.

The experiments reported here are centered around a situation which, it seems to us, should certainly have developed a specific tonicity in definite muscles which were to react, if such tonicity is a primary requirement of the set or orientational period.

Procedure

The reaction time situation was selected as being most adapted to the methods of detection employed. The records were of combined action potentials and gross movements. It was assumed that the warning period of a simple reaction response would be likely to fit closely into the general conception of set and orientation when the latter is used to designate the persisting tendency conditioning the appearance of a predetermined voluntary act.

The warning signal was the glow of a neon bulb, and the stimulus for the reaction was the disappearance of this light or the appearance of a glow in another nearby light.

At the appearance of the stimulus the reaction was usually an extension of the forearm about the elbow by the action of the triceps brachii. The results do not justify a detailed discussion of the variations in amount of training given the subject, and methods of giving the stimulus. Both choice and simple reactions were studied and in both types both motor and sensory sets were assumed by the subjects, a part of whom were trained in introspective method and should, therefore, have been able to adopt these sets successfully.

In all about 20 subjects were used, and over 150 reactions were photographed. In addition to these subjects a number of records were made of reactions in cases of hemiplegia, paralysis agitans and other organic types of disorder. These records were made in the hope that they might give some hint as to the processes of normal integration. To date no important facts have become apparent in these records, and they are, therefore, not discussed in this report.

The apparatus used can most conveniently be described under five headings:

(*a*) Apparatus for placing the subject in proper position to respond. This portion of the apparatus has been previously described by McKinley and Berkwitz (1928) and the necessary portion of that description is quoted. (Since the present apparatus was arranged for both arms instead of one as in the original, plurals have been added.)

"The patient is placed semi-recumbent in a dental chair with the trunk and head of the body at about a 60 degree angle from the horizontal. The height of the chair is so adjusted that the shoulders are higher than the arm rests of the machine and so that with the forearms and hands on the arm rests the upper arms form an angle of approximately 10 degrees with the horizontal. Variations up to 5 degrees in the angle of the arm from the horizontal appear to introduce no error in the curves. The chair is so placed in relation to the machine that the arms are as nearly as possible in the plane of abduction, that is, drawn neither anteriorly nor posteriorly. The latter position is important because retraction of the arm increases the tension on the biceps and decreases the tension on the triceps muscles, and a position too far anteriorly does just the reverse. The arm rests of the apparatus rotate freely on ball bearing spindles. . . ." To the above spindles were attached silk threads which led into the oscillograph box (see below) and which were attached to mirrors so that angular displacements of the arms would be recorded in a linear fashion by two spots of light on the film. Thus, it was possible for the subject, while in a relaxed position, to move his arms freely at the elbow, such movements being recorded in graphic fashion on the film as linear displacements of the recording light spots. The movements of the spots in response to movement of the subject's arm were delicate and hence afforded a suitable indication of the point at which the first movement of the arm occurred as well as recording the velocity and extensity of the response. Each spot was displaced about 2½ inches for a 90 degree change in the corresponding arm support. In the following experiments, the reaction

proper consisted in either the extension or the flexion of the forearm. In most cases, extension was used since it is more definitely dependent upon a single muscle, the triceps, than is flexion upon the biceps.

(*b*) The electrodes. As is usual in such experiments as these, the electrodes constituted a difficult problem. The main disadvantage in most of the electrodes experimented upon was their failure to isolate suitably the muscle to be investigated. Ordinary surface electrodes proved unsatisfactory. We next inserted two platinum wires into the muscle but found that this also failed to eliminate the effect of action potentials resulting from contractions of the antagonistic muscle. Our final development consisted of a modification of the pore electrode. We first used pore electrodes as described by Adrian and Bronk (1929), but found the output from these to be too small and too restricted. Our modification consisted in the development of a wider separation of the two electrodes and the control of this distance. This set-up also permitted one of the electrodes to have a relatively large area, and allowed the other a small area of contact, restricted definitely to an interior portion of the muscle under investigation. The neutral or outside electrode consisted of a gauge 19 hypodermic needle, 2 cm in length, which was sharpened to a blunter angle than usual. The active electrode consisted of a platinum iridium wire of .32 mm diameter, insulated by flexible baking varnish except for the sharpened tip; this was passed through the bore of the neutral electrode after the insertion of the latter into the muscle. The distance to which the active electrode projected beyond the neutral electrode was controlled by means of an insulated flange fitted into the large opening at the head of the grounded or neutral electrode. The exposed tip of the active electrode projected beyond the tip of the grounded one by about 4 mm. A terminal wire was soldered to the outer end of the active electrode; connection to the neutral electrode was effected by means of a clip applied to its head after insertion. For most of the experiments, one of these electrode assemblies was placed in the triceps and another in the biceps muscle of each arm of the subject. Once inserted, the electrode caused relatively little discomfort. Insertion was made with minimum disturbance or pain to the subject by the use of ethyl chloride applied locally. We also used for later experiments an electrode something like a thumb tack in shape. It had for the head or flat part a small 8 mm square of Monel metal through which projected a three cornered cutting needle of steel for about 5 mm. Two of these were placed over each muscle to be investigated at a distance of about 4 cm. The sharp tip entered the skin for 1 to 2 mm. These modified surface electrodes gave no results differing from the deep ones and offered a check upon possible unpleasant interferences in the deep type.

(*c*) The amplifiers. Four amplifiers were used, and since these were duplicated in circuit and parts, we shall describe only one. The circuits were of the usual resistance coupled type, having .1 micro-farad coupling condensers, and 100,000 ohm wire-wound plate resistances. The tubes used were one type PJ 11 and one type '22 as voltage amplifiers and one '47 pentode output tube. Control of amplification was secured through the use of a wire-wound potentiometer of 10,000 ohms which was arranged to govern the screen voltage of the '22 stage from 0 to 22.5 volts. An output transformer wound with a secondary matched for 15 ohms impedance was used on the pentode output tube, and an interstage coupling transformer was used as an input transformer. By careful selection of tubes, thermal effects and microphonics were cut to a minimum. The active leads to the amplifiers were carefully shielded. Since the amplifying characteristics of these instruments are of no importance independent of the recording apparatus, we shall omit discussion of this point until the completion of the next section.

(*d*) The recording apparatus. The recording apparatus used was a Westinghouse five-element oscillograph, in connection with the rotating drum camera and 3¼ inch wide 'Verichrome' kodak film.

Four of the oscillator vibrator elements were of the 'super-sensitive' type, consuming approximately 2 milliamperes per inch deflection with direct current, and the fifth element was of the 'standard' type consuming about .13 amperé per inch deflection on direct current. Four of the super-sensitive vibrators were used as recorders for the amplifier assemblies and the fifth was used in connection with the oscillator (see below) for the one-one hundredth second time record. This vibrator was also used as a signal. The adaptability of this type of oscillograph to such problems as those undertaken having been adequately discussed by Travis and Hunter (1928) does not require repetition.

(*e*) Auxiliary apparatus. No attempt was made to keep the amplifiers absolutely standardized for a given input-output ratio. They were, however, standardized approximately at the beginning and end of each experiment. For this purpose the output of a General Radio, low frequency oscillator, oscillating at 100 cycles per second was utilized. The amplifier inputs were shunted across one ohm, non-inductive resistance, and the current flowing through this shunt was read upon a 0.500 micro-ammeter. Thus, the micro-voltage on the input of the amplifier was approximately the micro-amperage flowing. Standardization was arbitrarily made at a frequency of 100 cycles per second. Though greater sensitivity was possible, no advantage was gained by using greater than .5 cm amplitude output to the camera for 50 micro-volts input. This sensitivity was accordingly used for most of the data. It should be kept in mind that at the greatest sensitivity used, the base line was quite regular and quiet when the subject was relaxed. The frequency response of the amplifiers was good, with a maximum of efficiency in the neighborhood of 100 cycles and with relatively little cut-off up to one thousand cycles. Frequencies as low as 25 cycles were reproduced with little distortion though there was a loss of efficiency. It is important to remember that micro-voltage readings of action potentials are always complicated by the presence of surrounding tissue which, of course, shorts away a considerable per cent of the actual potential. In general, the nearer the active electrode lies to the acting tissue the greater will be the recorded potentials. The stimulus lights were controlled by a commutator on the oscillograph driving head.

Results

Figure 1 is a typical record of a choice reaction in the left arm. The alternative reaction was an extension of the right arm. In this the time line is the signal line showing a break downward when the warning light goes on and an additional drop for the stimulus light. The vibrations of this time line are 10σ. Arm movement lines are directly below the time line. The wide one is that of the left arm and the narrow one that of the right arm. Each deviates upward on extension of the corresponding arm. The lower two amplifier lines are those for the left arm and the upper two those for the right

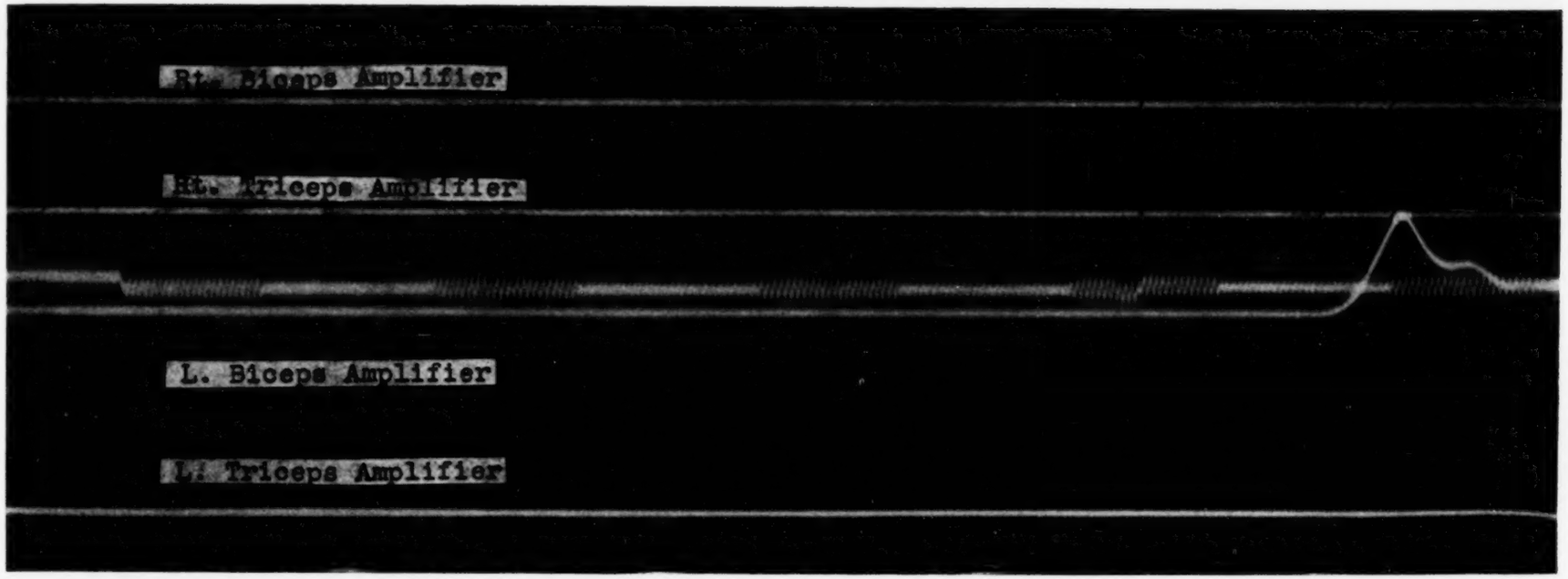

FIG. 1. A choice reaction in which the response was an extension of the left arm. The movement of the right arm is probably due to gross movement of the body caused by the sudden extension of the left arm. The warning signal is near the left end of the record and the warning period is about 1.5 seconds.

arm. In each pair the lower one records from the triceps and the upper the biceps. Therefore, in reactions involving extension the active muscle will be the lower of the two. Variations in the signal line will be noted to appear occasionally, through induction in the four amplifier records, as variations of a quite constant nature. These should be disregarded. It will be seen by examination of the action potential curves on this record that the typical onset of these curves is somewhat gradual, rising within about 20σ to a large variation. It is typical of most of the records that both the triceps and biceps begin to react practically simultaneously irrespective of which is actually to make the reaction, although in most cases the muscle which moves the arm in a given direction shows the largest action potential. The effect is probably not due to carry-over of one amplifier into the other as may be more clearly indicated from an examination of the curves for the two muscles in Fig. 2. It will be noted that they show practically independent activity though both are involved. In further confirmation of this fact, several records were made in which every precaution was used to allow the subject to contract only one of the two muscles involved. These records were negative with reference to carry-over.

The instructions to the subject were that he regard the appearance of the first light as a warning, and react as quickly as possible with extension of the arm at the appearance of the stimulus light. For choice reactions two stimulus lights were used and the subject was usually instructed to react with the left arm to the right light or vice versa. A number of records were taken in which flexion of the arm was the response, but these need not be considered separately. Instructions varied from subject to subject in attempts to develop instructions which would give more definite preparatory sets, but the final results do not show any important differentiation for the different kinds of instructions.

Records were kept of the subjects' introspective reports of the reactions, and an attempt was made to correlate the presence of particular reports with particular types of reaction curves. This correspondence was not observed although two

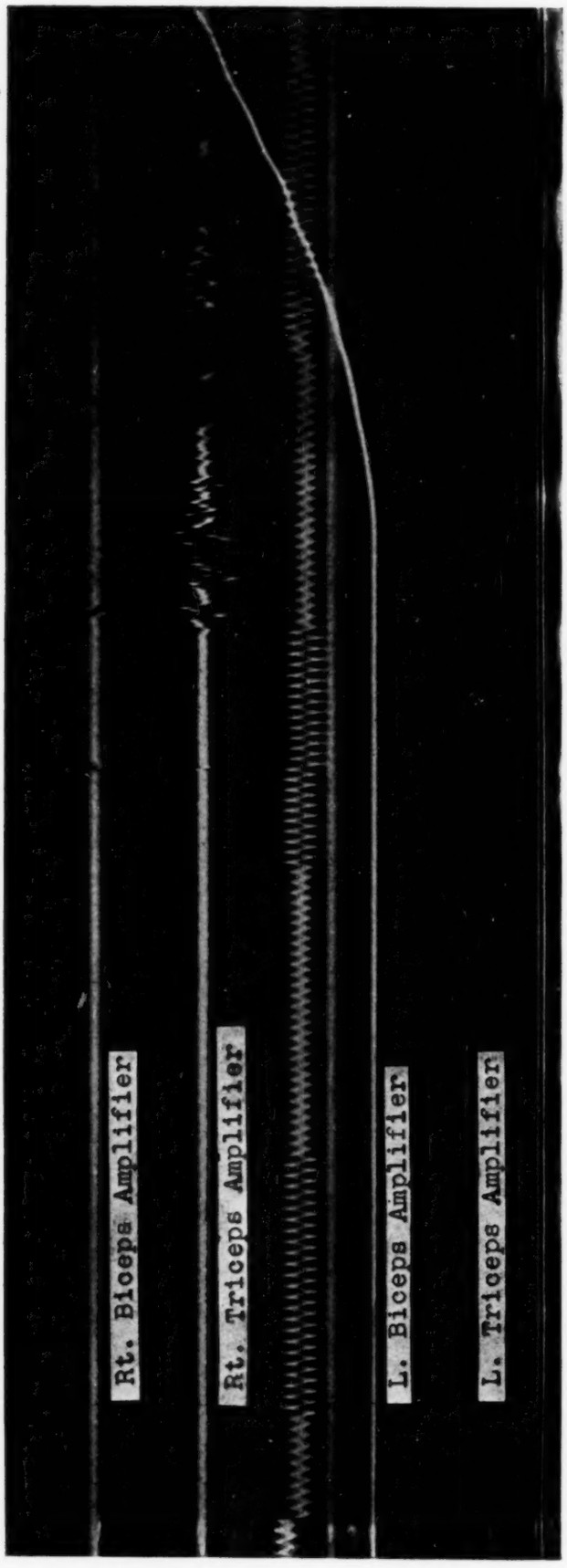

FIG. 2. A moderate speed extension of the right arm showing the involvement of the antagonistic muscle.

of the subjects were well trained in introspection. There were, of course, the usual variations in reaction time. All of the subjects reported a general feeling of muscular readiness for the stimulus and a particularly definite increase of tension as the warning time was made longer preceding the stimulus. In some cases, about 25 trials were given in practice before the taking of any records, although the amount of practice appeared to have no essential effect upon the final expression of the reaction. Many of the records were taken immediately following a series of non-recorded reactions in which the warning time was considerably shorter than that of the recorded reaction. This procedure was adopted with the idea that the longer warning time occurring in the actual record should be the best condition for the development of a possible tension in the particular muscle which was to bring about the reaction. The introspective reports of the subject indicated that this increase of tension with the increased length of the warning period really occurred in the case of the reactions photographed. The actual warning period varied in length from time to time, but this apparently had no effect upon the reaction as expressed in the curves.

In spite of this deliberate effort to develop such specific tension, we have not a single record which shows such tension developing within the warning period. Ten out of fifty-five records show intermittent tension in the muscle to be used in the response; this tension was present at the first of the picture and was apparently not influenced by the warning signal. Three additional records indicate this tension in the antagonistic muscle.

Our results then, are negative with regard to the necessity of specific tension in the muscle which is to do the responding. Whether or not it is possible to develop this type of tension by some other procedures than those used is, of course, not definitely settled. We feel, however, that the results, as they are exhibited in these records, indicate that the presence of these tensions is not as frequent as might be expected from a motor psychology based upon a completed reflex arc.

There was a delay of varying amount between the first

appearance of potentials in the amplifier records and the first detectable movement of the arm. This may be due in part to a small amount of inertia in the arm movement recording device. This delay appears, however, to vary between a minimum of about 60σ and a maximum of 120σ or more, depending upon the force in the reaction.

In order further to study the involvement of the antagonistic muscle a number of records were made in which the subject was instructed to make the reaction more slowly. The speeds of arm movement in these records varied from very slow to the maximum he could accomplish. Figure 2 is typical of one of these where the speed of movement was moderate. In general the antagonistic muscle is slightly active in all reactions, even very slow ones. There is some evidence that in the very slowest movement it is more active than in those of slightly greater speed. This may be a tensing to provide more control for the extremely slow movement.

In general there is no involvement of the arm opposite to the responding one. In choice reactions (Fig. 1) there may be a slight contraction but this is not constant.

In our subjects a further modification of the response pattern was studied. These subjects were directed to extend both arms upon presentation of the stimulus. Records were taken at high speeds. One of these is illustrated in Fig. 3. There was some constancy in regard to the order of appearance of action potentials in the four muscles observed. This constancy could not, however, be correlated with handedness or any other factor. We are continuing these experiments hoping to find some significance in the findings. Most constancy was in the appearance of activity in the extensors or acting muscles several sigma before that in the antagonistic. In some records, however, this difference is doubtful.

Discussion

Our failure to demonstrate the invariable presence during the warning period of a preliminary tone or tonus increase in the muscle leads naturally to a consideration of the sensitivity of the apparatus and of the electrodes for picking up any

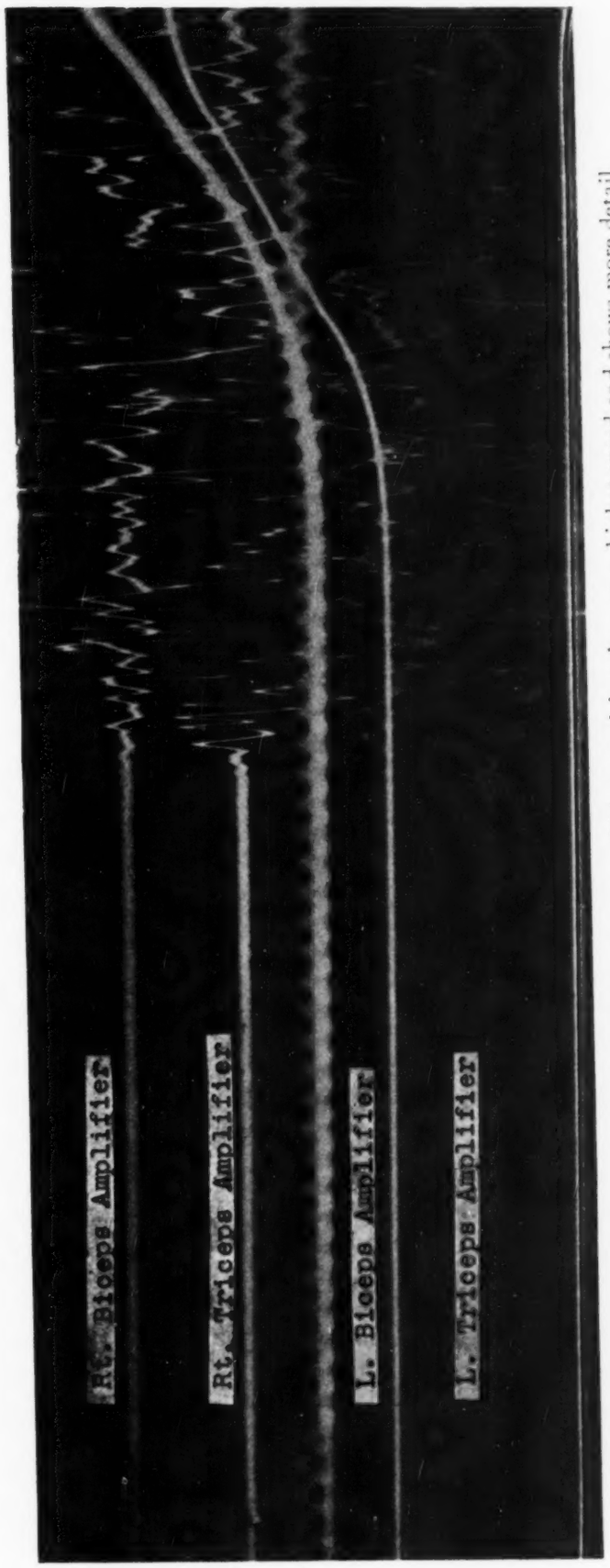

FIG. 3. A simultaneous extension of both arms. This record is taken at a higher speed and shows more detail.

changes which might be there. As has been shown in the description of the apparatus and electrodes, the sensitivity is high. Considering the amplifier and recording system solely, the sensitivity approaches the utmost which is possible at present with electron tube amplification where the base line is to remain steady. In that connection it would have been possible for us to increase the sensitivity, but it would have been at the expense of the quiet base line, and an allowance would have to have been made for every response which was observed, depending upon the normal variation in the unsteady base line obtained in that manner. It seemed to us better to avoid such correction and to adopt rather a somewhat lower sensitivity with the expectation of being able to rely upon variations from that base line as definitely indicating the presence of an action potential. An indication that we obtained this desired result is given in control records which we have taken in which the subject simply sat completely relaxed and did not respond in any way during the picture. These records invariably show a quiet base line.

From the foregoing it seems highly probable that if there were potential changes in the muscle, we would have obtained in a few cases at least, some evidence of those changes. It would then be indicated that if such preliminary tightening of the muscle occurs as is implied in the terms 'postural substrate' or 'associative movements' or in any system which requires that the specific muscle must be partially contracted in order that a set for reaction may be present in the individual, then we must say that either such contractions do not exhibit action potentials or that they are not of necessity present in the specific muscle or in the antagonistic muscle involved in the reaction. This does not eliminate the possibilities that tension of various sorts may be present in the throat, chest or other musculature. If such were the fact, it would point to the conclusion that such tensions are of an extremely flexible nature or that the real set or orientation is independent of these postural changes. Indeed, it seems more logical to regard such postural changes as analogous to a string about the finger as a reminder to do something. The

string in itself is not an orienting mechanism; it is only an evidence of the existence back of it of the orienting tendency and does not give evidence as to what tendency is present. The shifting of the tension from the forearm to perhaps the chest, or neck or eyes in the same subject as reported introspectively, offers further evidence for the secondary nature of these tensions. This evidence does not rule out the evidence of Jacobson that during mental activities there may be changes in tension of specific muscles. The evidence does indicate that such tensions are not necessarily present in the specific muscles. In this connection it should be noted that in Jacobson's experiment there was no control of the degree or vividness with which the subject imagined the particular movement examined. It may be true that it is not possible to imagine waving the arm without actually using some eye, throat, or arm muscles. It may also be true that these are simply an indication of cortical activity and that the same contractions may result from very widely differing mental activities.

The simultaneous appearance of action potentials in the two antagonistic muscles has been previously observed. One fact should be kept in mind and that is that experiments upon the relaxation of the antagonistic muscle have been conducted in most cases upon either decerebrate or otherwise mutilated animals. It may be that the presence of the cerebrum is a complicating factor to the sequence of what otherwise would be well organized reflex activity. This point has already been suggested by several investigators (Tilney and Pike, 1925). The evidence present is obtained differently but is corroborative.

The antagonistic muscle is more and more involved in the reaction in direct proportion to the acting muscle. A part of this activity may be serving as a buffer to the throw of the arm, but if it were all to be explained in this way, there should be no activity for slow movements until the movement had got under way. Contrary to this, activity appears in all cases before any gross movement of the arm. These results suggest once again that great caution be exercised in interpretation of higher coördinations in terms of simple reflexes.

Summary

Within the limits of sensitivity of the apparatus, we find no evidence for the appearance of increased muscular tension in either the muscle which is being used for the response or in its antagonist during the warning period preceding a simple or choice reaction.

The actual response is characterized by the nearly simultaneous appearance of action potentials in the responding muscle and in its antagonist. These action potentials first make their appearance at a minimum of about 60σ before the appearance of actual movement in the arm. The two or three preliminary small discharges are quickly followed by larger ones which may be predominantly large in the muscle effecting the response.

(Manuscript received May 19, 1934)

References

1. Adrian, E. D. and Bronk, D. W., The discharge of impulses in motor nerve fibers. II. The frequency of discharge in reflex and voluntary contractions, *J. of Physiol.*, **67**, 119–151, 1929.
2. Freeman, G. L., The spread of neuromuscular activity during mental work, *J. of Genetic Psychol.*, **5**, 479–494, 1931.
3. Jacobson, E., Electrical measurements of neuromuscular states during mental activities. I. Imagination of movement involving skeletal muscle, *Am. J. of Physiol.*, **91**, 567–608, 1930*a*.
4. ——, Electrical measurements of neuromuscular states during mental activities. IV. Evidence of contraction of specific muscles during imagination, *Am. J. Physiol.*, **95**, 703–712, 1930*b*.
5. ——, Electrical measurements of neuromuscular states during mental activities. V. Variation in specific muscles contracting during imagination, *Am. J. Physiol.*, **96**, 115–121, 1931.
6. McKinley, J. C., and Berkwitz, N. J., Quantitative studies on human muscle tonus, *Arch. of Neur. & Psychiatry*, **19**, 1036–1056, 1928.
7. Tilney. F., and Pike, F. H., Muscular coördination experimentally studied in its relation to the cerebellum, *Arch. Neur. & Psychiatry*, **13**, 289–334, 1925.
8. Travis, L. E., and Hunter, T. A., The relation between 'intelligence' and reflex conduction rate, J. Exp. Psychol., **11**, 342–354.

THE CONDITIONS OF VERBAL CONFIGURATION

BY H. L. HOLLINGWORTH

Columbia University

Advocates of the gestalt conception have pointed out that a configuration does not arise equally easily under all conditions. Thus we are told that the ape sees a branch more readily as a tool for reaching when it is not attached to a tree but is instead detached from this structure and lying 'loose' on the ground. And the problem situation 'closes' upon the stick as such an implement more easily if the stick lies between the animal and the goal, or if stick and goal are close together in visual space. Similar observations have been made in the case of young children, and in the case of idiots.

Now words would seem to be configurations *par excellence.* It was early shown that words are identified as *wholes* or as *forms* about as easily as are their component letters. The total form may *carry meaning* even when the separate letters are not all clearly discerned. Indeed, missing letters are *supplied*, which would appear to be a neat example of 'closure.' And wrong letters are mis-read, errors in spelling overlooked, which would seem to illustrate the contention that it is the whole that determines the part.

If words are typical configurations, as they appear to be, we might then expect that word building would be a process which would neatly exemplify the various laws of configuration, and the conditions under which configuration most readily occurs. In the following experiment we have used word building experimentally, in order to see whether these expectations are to be realized.

Arguing from analogy, we might expect that words could be 'built' from letters more readily or rapidly if these letters were presented mixed up in such a way that they did not already constitute a 'word.' Thus it ought to be easier to construct the word 'intellectual' from the letters 'alinletluetc'

than from the letters 'all in lettuce,' in spite of the fact that the letters are identical in the two cases. In the second case, however, they already participate in a 'configuration,' and according to the law it ought to be more difficult to wrest them from this organization and put them together differently, than it would be in the first case, in which the combination does not *mean* anything to start with.

And if such a result should be found, it might be reasonable to expect that individuals would differ in this respect, and that such differences would constitute important features of mental make up. They might even, if found, be related to the casual observations that some people find it harder than others to abandon an idea which they have once espoused. And such tenacity, or adhesion, or refractoriness might be sufficiently important to be given a technical name, and to have diagnostic tests constructed for its identification and measurement.

Our first experiment was as follows. A brief and fairly interesting short story, of one typewritten page, was written, the whole experiment having been suggested by a 'parlor game' that was perpetrated on the writer at a social function. Of the words of this story twenty-five were 'disarranged'; the letters were all there but by the jumbling of the letters the 'word' was concealed. Two forms, *A* and *B*, were prepared, in which these same twenty-five words were 'disarranged.' In form *A* the jumbled letters did not 'make sense'; thus 'intellectual' appeared in the form of "alinletluetc"; in form *B* this word appeared in the form 'all in lettuce' (more strictly 'allinlettuce').

College students were given a fixed time in which to *correct* the story, writing above each jumbled word the letters in their correct form. Half the group were given form *A* and the others form *B*. When 50 cases had been secured for each form, the papers were scored for the number of words correctly identified, and a very striking difference was found. An attempt was made to verify this finding on two more groups of 50 subjects each, and an equally striking difference was found, but in just the opposite direction. The average

result was just no difference between the two forms. Addition of further cases failed to produce any reliable difference.

We then sought to simplify the conditions, by eliminating the story, with its distractions and contextual suggestions. The new material consisted of a list of 50 names of familiar animals, such as tadpole, oyster, grasshopper, etc. Again two forms were prepared. In form *A* the letters were disarranged in such a fashion that they not only made no familiar animal name, but did not constitute a word at all. In form *B* the letters were disarranged in such a way that they formed words or word pairs, but not the names of animals. Thus in form *A* the animal 'tadpole' appeared as 'lpae odt'; in form *B* this animal appeared as 'leap dot.'

Various other rules were applied, such as keeping, so far as possible, a given letter in the same part of the total combination in both cases; breaking the meaningless letters into two parts if the meaningful arrangement took the form of two words; etc. The instructions given were as follows:

INSTRUCTIONS

All these are the names of animals, with the letters disarranged. Thus of CUB KOLL you can make the word BULLOCK; and of MECLA you can make the word CAMEL, and so on.

Opposite each write the correct name of the animal, using all the letters given, and no others. Do as many as you can in the time allowed. If some seem difficult at first pass over them and come back to them later, if you have time. Time will be called in ten minutes from the beginning.

Neither form *A* nor form *B*, as described, was used. Instead derived forms *C* and *D* were prepared. Each contained the whole 50 names or arrangements, half from form *A* and half from form *B*, in regular alternation, thus, in form *C*,

lpae odt	(tadpole)
storey	(oyster)
gpoerh psrsa	(grasshopper)
rag bed	(badger)

In form *D* the word appearing in one way in form *C* was here shown in the opposite fashion. Thus form *D* began as follows.

leap dot	(tadpole)
soetyr	(oyster)
gopher spar	(grasshopper)
rga bde	(badger)

The correct word formations are given above in parenthesis, although of course these did not appear on the test blank. The two Forms would thus, it was hoped, serve as mutual controls, and could be combined in the final figures. Each contained the names of all the animals, half in one way and half in the other, the two simply being reversed in this respect. Any given individual was given but one form.

From the results various scores could be secured. In the first place, the total number of correct words built by each subject could be determined. These scores ranged from 1 to 25. Such scores do not bear directly on our problem, although they may tell something interesting about the individual. The process of word building deserves more careful analysis and study than it seems to have been accorded.

The only thing we have done with these scores is to compare them with Alpha Test intelligence scores which were secured in case of the majority of the subjects, of whom there were about 75 in each group (form *C* and form *D*). The results are as follows.

Median Words Built	Alpha Score					
	Over 180	170–179	160–169	150–159	140–149	Below 140
Form *C*.....	11.5	10	11	11	8	11
Form *D*.....	10.5	6.5	6.5	8	5.5	5

There is no consistent relation between Alpha score and the number of words built, among these college students, all of whom had already a certain familiarity with various forms of mental tests, through having been submitted to them. The range of Alpha scores, to be sure is small; but except for one group in form *D*, those scoring high in Alpha appear to build no more words than those with low scores. The only clear result that does appear is that form *D* is apparently more difficult than form *C*.

A more relevant score is to be derived from the number of words built from 'configured' and from 'unconfigured' letters in the case of each individual. The same difficulty arises here that arises always in the discussion of configuration. Of course by the absence of configuration is scarcely meant that the features have *absolutely no arrangement.* Just being there, side by side is of course an arrangement. Perhaps we can take refuge in the gestaltist's term 'prägnanz,' and call the arrangement which is more meaningful, that is which constitutes *some* word, a more *pregnant* configuration than that which does not do so, even though the *some word* is not the name of an animal.

In form *C* the average number of words (73 cases) built from configured letters was 5.4; from unconfigured letters the average was 5.7. In form *D* the average number of words built (79 cases) from configured letters was 3.7; from unconfigured letters the average was 4.5.

This looks promising; in both cases more words are built from letters that do not already constitute a configuration. The college student resembles the ape, the infant, and the idiot.

Unfortunately for the neatness of this outcome, we now secured another group of 42 cases on form *C.* The average from configured letters was 4.7; from unconfigured letters the average was 2.9. This contradicts not only the result from form *D* but also the previous result from form *C.* Combining all the available cases, form *C* (115 cases) and form *D* (79 cases) give opposite results, and no conclusion whatever can be drawn from these scores, as averages.

We next treated each subject individually, inquiring whether he built more words from one arrangement of letters than from the other (configured or unconfigured). We took the number made from unconfigured words as unity and stated each score by dividing this number into the number built by that individual from configured letters. A ratio of 100 thus means no difference; ratios greater than 100 mean more words built of configured letters; ratios less than 100 mean more words built of unconfigured letters.

The median ratio for form *C* is 100, while for form *D* it is 80 (73 and 79 cases respectively). The median ratio for the extra group of 42 cases on form *C* is also 100. Again the results are inconclusive. Form *C* would lead us to suspect just no difference. Form *D* would suggest that words are more easily built when the letters are 'free floating,' thus supporting the gestaltist expectation. The net outcome is at least not very favorable to the conception of a word as a constraining gestalt.

Again the subjects were distributed according to score in the Alpha Test, and the median ratios for each group found. The results are as follows.

Median Ratios	Alpha Scores					
	Over 180	170 to 179	160 to 169	150 to 159	140 to 149	Below 140
Form *C*	118	100	100	60	100	88
Form *D*	92	100	55	112	90	66
Average	105	100	78	86	95	77

The general trend of these medians suggests that the ratio is inclined to decrease with decreasing intelligence scores. The facts are somewhat clearer if we group the cases more roughly, into three groups instead of into six, as follows.

	Alpha Scores		
	Over 170	150 to 170	Below 150
Form *C*	109	80	94
Cases	38	33	18
Form *D*	96	84	77
Cases	26	14	9
Average	—	—	—

The figures are not very reliable, but it seems to be clear enough that those who are seriously handicapped by having the letters presented in a configured fashion are the students with Alpha scores below 170, these comprising about half of the group. The more intelligent half do not suffer this handi-

cap; if anything they profit slightly from it (average ratio 103). The presence of prepared configurations does not interfere with their wresting the letters apart and putting them into new forms.

As a matter of fact, even this attempt to get configuration into the character of a word is specious, and only proceeds by ignoring the unreliability of the differences. If we group together all those having ratios over 100, all those having ratios of exactly 100, and all those with ratios below 100, the median Alpha scores for the three groups are 166, 170 and 161 respectively. Here the highest scores actually belong to those who show just no effect of the configuration.

Conclusions

It seems necessary to conclude that we have been unable to show that the word, as a configuration, exercises any constraining influence over the letters which compose it, at least in the case of subjects capable of making Alpha scores of 140 or over. Since the evidence suggesting a constraining effect was slightly greater with the group making lower Alpha scores, it perhaps remains an open question whether positive results might not be secured with duller or younger subjects.

As a matter of fact the results are just what we ought to have expected, on systematic grounds. For one arrangement of letters is just as much of a 'configuration' as is any other, in spite of the fact that a naïve person may feel that letters that make a word somehow 'look more like a unit' than do letters that make non-sense.

Tachistoscopic studies long ago demonstrated that words do not function as 'gestalten.' The whole word has meaning just in so far as it leads to consequents appropriate to previous 'occasions' of which it has been a part, as Burke long ago showed, in his 'Essay on the Sublime and the Beautiful.'

Parts of words may have the same effectiveness as whole words, this being only the familiar law of cue reduction. The 'shape' of the word, certain letters of the word, etc., may lead to writing or speaking or other activity just as effectively as the whole word would if it were there. This is only the ordi-

nary law of perception. Therefore when *parts* of words are exposed in a tachistoscope the subject, as is to be expected, reports a *whole* word.

But this is not an incomplete gestalt, exhibiting 'closure.' Nothing closes; and nothing is 'supplemented'; nothing 'completes itself.' The simple fact is that the *part* of the word functions as effectively as the whole word would, *were it presented instead.* Hence the term 'gestalt,' as applied to words, is only a smoke screen. The simple fact is that partial stimuli may function redintegratively. This does not mean, as in the old days of Sir William Hamilton, that *the rest* of the word appears in the visual field (although it may appear in the subject's report, which is another matter).

We are disposed to carry the conclusion over to all the other cases in which closure, completion, supplementation have been introduced as explanatory terms. Words, we have concluded, are not 'gestalten.' But they come as near being this as does anything else. When a figure is presented on the retina, part of it coinciding with the normal blind spot, or with some other blind area, we should expect the subject to 'reproduce' a complete figure, within the limits of his experience. This does not mean that the retinal pattern, or any pattern of 'brain stress' has 'completed itself.' It means only that the subject gives the same report that he would give had the whole figure which he now draws been presented.

People are indisposed to admit that they do not see every thing that is put before them, or that goes on around them. On the witness stand they 'patch out' the memory gaps by what seems to them to be reasonable material. In the laboratory experiment they do the same thing. Partial stimuli evoke reports and reactions appropriate to larger wholes. But to suppose that the wholes are themselves evoked or educed is only the gestaltist's illusion.

(Manuscript received April 7, 1934)

AN ATTEMPT TO CONDITION HAND WITHDRAWAL RESPONSES IN HUMAN SUBJECTS

BY HULSEY CASON

University of Wisconsin.

Introduction

The conditioned stimulus in the present study was a sound, the conditioning stimulus was an electric shock applied to the end of the right forefinger, and the original response was a withdrawal movement of the hand. Graduate students in psychology acted as subjects, and all were familiar with the topic being investigated. The experiment was carried out at Columbia University in March and April, 1922, and a written report was prepared at that time. This report was not published sooner because it is only recently that the experiment has seemed to have any value at all. According to Razran's review of conditioned withdrawal responses with shock as the conditioning stimulus in adult human subjects (**11**), 16 studies of this topic were made before 1922, and 12 of these were reported in the Russian language. Thirty one studies have been made since 1922, and if Ginsberg's study (**7**) were included the total since 1922 would be at least 32. Other studies of conditioned hand withdrawal responses with children as subjects have also been made (**10**).

When the present experiment was carried out I was especially interested in investigations of Bechtereff (**1**, pp. 181–186, 249–267), Watson (**13**), and Hamel (**8**), and was trying to determine the maximum speed of conditioned hand withdrawal responses. I hoped to obtain some data on the role of voluntary factors in conditioned motor responses of human subjects, and also hoped to test some of the claims for the conditioned reflex as a new psychological method.

General Procedure and Apparatus

We first tried to persuade each *S* to take as much of the electric shock as he could or would stand, and the shock used in the experiment was considerably stronger than this amount. *S* withdrew his hand without any hesitation every time the shock was turned on. From time to time during most of the experiments I would try to persuade *S* to take a stronger shock.

The conditioned stimuli used were the sounds of (1) a telegraph sounder, (2) a telegraph relay, and (3) an electric bell. The apparatus and electric connections were such that we could give (*A*) sound alone, (*B*) shock alone, (*C*) sound and shock simultaneously, or (*D*) sound .010, .020, or .040 sec before the shock. The time which elapsed between the sound stimulus and a hand withdrawal movement, if a withdrawal movement occurred, was in all cases measured in .001 sec.

The apparatus used in giving the sound stimuli and in measuring the time of the hand withdrawal responses was practically the same as the apparatus which was used in another experiment (3, 4). All time measurements were made on a Bergström chronoscope, and the accuracy of this instrument was checked a number of times against a properly adjusted Hipp chronoscope (3, pp. 161–162). The latent time of the telegraph relays and the sound producing apparatus was taken into consideration (4).

The general procedure used in most of the 12 experiments was as follows.

1. A number of measures were made of *S*'s fastest voluntary reaction time (V. R. T.) to the sound stimulus. No ready signals were used on any occasion in any experiment.

2. Sound and shock were the training stimuli in all experiments, and these stimuli were repeated a large number of times during the training period. They were given simultaneously or the sound was given a short time before the shock. The training stimuli were given irregularly at an average rate of about four repetitions a minute. There was a rest period of some four minutes after every 30 repetitions of the training stimuli. In each repetition the time which elapsed between the sound stimulus and the withdrawal response was measured in .001 sec, or σ.

3. After a variable number of repetitions of the training stimuli, but generally after 10 or 20, the sound was given alone and the time of any hand withdrawal (conditioned) response that occurred was measured.

4. For unconditioning, the sound stimulus was then given alone a number of times, and the time of the responses that occurred was measured.

5. Finally, a number of measures were again made of V. R. T. to the same sound stimulus.

The individual experiment for each *S* lasted between two and three hours.

Results for the First Four Subjects

In the case of the first four *S*'s, the right hand rested on a broad metal electrode and the end of the right forefinger touched a metal rod which was placed in a convenient position. The metal rod served as the active electrode. A wooden rod 1.5 cm in diameter was immediately above and just touching the third phalanx of the *S*'s right forefinger. When *S* jerked his finger away, it touched the wooden rod and per-

manently closed an electric circuit which stopped the chronoscope. The time which elapsed between the sound (and shock), the sound alone, or the shock alone, and the moment *S* touched the wooden rod, was measured in each reaction. The procedure and results for each of the first four *S*'s were as follows.

Experiment 1, Subject Mr. J. H.—Training stimuli: telegraph sounder and shock given simultaneously 323 times, and median time of response to both stimuli was 180.0 ± 1.1σ. Sound alone given after irregular intervals 37 times during training period. Withdrawal responses (C–R's) occurred in 34 trials, and median time of these 34 C–R's was 183.0 ± 2.9σ. For unconditioning, the sounder was given alone 36 times, a hand withdrawal response occurred 21 times in trials 1–13, 15, 18, 20, 22, 24, 25, 26, and 30, and the median time of these responses was 189.0 ± 11.8σ. Fifteen measures were then made of V. R. T. with the same sound stimulus and the same withdrawal movement, and the median time was 252.0 ± 27.6σ. This *S* seemed to be reacting voluntarily on all occasions.

Exp. 2, Mr. B. C.—Forty-two measures made of V. R. T. to telegraph sounder, median time 191.5 ± 6.3σ. Training stimuli: sounder and shock given simultaneously 335 times, median time of response to both stimuli 166.0 ± 1.2σ. Sound alone or shock alone given after irregular intervals during training period, each 22 times, and a response occurred on all trials; median C–R time to sound alone 169.0 ± 5.7σ, and median time of responses to shock alone 168.0 ± 3.1σ. Forty-two measures then made of V. R. T. with same stimulus and same response, median time 198.5 ± 8.4σ. After the experiment had been completed, *S* said that when he withdrew his hand to the sound alone, he felt that he had been fooled into reacting and that he had acted like a mechanism. When the shock was given alone he felt that he should not have jerked his hand away and that he had done something that was not required. The median times of responses to sound and shock (166σ), to sound alone (169σ), and to shock alone (168σ) were practically the same.

Exp. 3, *Miss E. B.*—Training stimuli: sound of telegraph relay and shock given simultaneously 163 times, median time of response to both stimuli 174.0 ± 1.7σ. Sound alone given 17 times during training period, C–R's occurred in 15 trials, median time 188.0 ± 12.8σ. For unconditioning, sound then given alone 60 times, responses occurred 45 times in trials 1–8, 12–31, 39–48, 50–53, and 55–57, median time of these 45 responses 194.0 ± 6.3σ. Finally 26 measures made of V. R. T., median time 273.5 ± 17.4σ. One peculiarity of this experiment was that C–R time (188σ) and time of responses in unconditioning (194σ) were practically the same.

Exp. 4, *Miss H. S.*—Twenty-five measures made of V. R. T. to sound of telegraph relay, median time 306.0 ± 6.2σ. Training stimuli: sound of relay and shock given simultaneously 220 times, median time of response to both stimuli 232.0 ± 1.8σ. Sound given alone 25 times during training period, C–R's occurred in 23 trials, median time 245.0 ± 18.4σ For unconditioning, sound then given alone 19 times, responses occurred only on first 10 trials, median time 298.5 ± 28.3σ. Finally, 25 measures made of V. R. T., median time 279.0 ± 8.7σ. The *S*'s whole body jerked every time the shock was turned on.

Improvements in the Apparatus

In the four experiments described above, *E* and *S* were in the same room separated by a screen, and in the remaining experiments, *E* and *S* were in different rooms and *S* could not hear the apparatus in *E*'s room. In the first four experiments, *S* had to give some attention to the position of his right forefinger, and the time which elapsed before his forefinger struck the wooden rod was somewhat variable. For the experiments described below, *S* kept the end of his right forefinger on a metal disc 1.1 cm. in diameter, and the disc was mounted on the end of an inverted telegraph key. The metal disc served as the active electrode, and a minimum pressure of 90 gr. was necessary to break the connection in the key. *S* kept two fingers of his left hand in a weak salt solution which served as the other electrode. When the shock was turned on, *S* released the key and this closed an electric circuit and stopped the chronoscope. The exact position of *S*'s right hand was not now of any particular importance. Most of the other details of the apparatus and procedure were the same as in the first four experiments.

Results for the Next Three Subjects

Exp. 5, *Mr. L. G.*—Forty measures made of V. R. T. to telegraph relay, median time 185.0 ± 3.8σ. Training stimuli:

sound of relay and shock given simultaneously 323 times, median time of response to both stimuli 143.0 ± 0.6σ. Sound given alone 37 times during training period, C–R's occurred in 36 trials, median time 157.0 ± 2.2σ. For unconditioning, sound then given alone 13 times, responses occurred only on first four trials, and the individual times were 201, 267, 590, and 435σ. Finally, 40 measures made of V. R. T., median time 215.0 ± 5.0σ.

Exp. 6, Mr. L. G.—*S* reacted voluntarily to a telegraph sounder in every reaction. This experiment was not directly concerned with the conditioned response, but with the influence of electric shock and the factor of expectancy on V. R. T. Mr. L. G. was the *S* in both the present and preceding experiments, and he was a practised *S* in reaction time experiments. The sound and shock were the same in both of these experiments. The present experiment may be divided into 11 consecutive series of 30 reactions each. In the 1st, 7th, 9th, and 11th series, *S* was assured that he would not receive a shock, and while he was not expecting a shock, 120 measures were made of speed of response to sound alone, median time 153.5 ± 2.5σ. In the 2d to 6th, 8th, and 10th series, *S* was informed that the shock would be given simultaneously with the sound in an irregular way half of the time, and while he was expecting the shock to be given half of the time, 105 measures were made of speed of response to sound alone, median time 136.0 ± 5.2σ. While *S* was expecting the shock to be given half of the time, 105 measures were also made of speed of response to both sound and shock, median time 119.5 ± 1.2σ.

It is possible to make some comparisons between the results obtained in these two experiments. In Exp. 5, C–R time to sound alone (157σ) was 28σ faster than V. R. T. (185σ, measured at beginning of experiment). In Exp. 6, V. R. T. to sound alone was 154σ when not expecting the shock, and 136σ when expecting the shock to be given half of the time, showing that expecting the shock decreased the time 18σ. If these 18σ be subtracted from the difference of 28σ (between C–R time and first V. R. T. in Exp. 5) in favor of a C–R, it would leave 10σ in favor of conditioning in Exp. 5;

but these comparisons are not entirely valid. While expecting the shock in Exp. 6, the direct stimulating effect of the shock on speed of response was 136–120, or 16σ, and the direct stimulating effect of the shock in Exp. 5 was 157–143, or 14σ. The presence of the shock had approximately the same influence on speed of response regardless of whether *S* was under special instructions to react voluntarily to the sound or not. When the shock was not expected, the median V. R. T. to sound alone was 154σ in Exp. 6 and 185σ and 215σ in Exp. 5. V. R. T. was 154σ when not expecting the shock in Exp. 6, but C–R time was 157σ when expecting the shock in Exp. 5.

Exp. 7, Mr. J. W.—Forty measures made of V. R. T. to sound of telegraph relay, median time 218.5 $\pm$ 5.7σ. Training stimuli: relay and shock given simultaneously 135 times, median time of response to both stimuli 145.0 $\pm$ 0.8σ. Sound given alone 15 times during training period, C–R's occurred only on first and fourth trials, these individual times being 205 and 308σ. Finally, 40 measures made of V. R. T., median time 199.0 $\pm$ 4.3σ.

Successive Conditioned Responses

In the training period for the following five *S*'s the conditioned sound stimulus was given 10, 20, or 40σ before the conditioning shock stimulus. A simple pendulum-contact device was used to give the sound 10, 20, or 40σ before the shock in each repetition of the training stimuli. In each reaction, a time measure was obtained of the interval between the sound stimulus and the hand withdrawal movement.

Exp. 8, Mr. A. R.—Sixty measures made of V. R. T. to telegraph sounder, median time 169.0 $\pm$ 5.0σ. Training stimuli: sounder and shock, sound given 10σ before shock, stimuli repeated 223 times, median time of response to both stimuli 145.0 $\pm$ 1.5σ. Sound given alone 17 times during training period, C–R's occurred in all trials, median time 188.0 $\pm$ 12.4σ. Finally, 30 measures made of V. R. T., median time 225.0 $\pm$ 6.9σ. After the experiment had been completed, *S* said that when the sound alone produced a C–R he did not think about withdrawing his finger and was only

aware of its being off the key after it was off. The sound and shock seemed to come at the same time during the training period.

Exp. 9, Mr. L. A.—Fifty measures made of V. R. T. to very loud telegraph sounder, median time 135.0 ± 1.8σ. Training stimuli: very loud sounder and shock, sound given 20σ before shock, stimuli repeated 160 times, median time of response to both stimuli 129.0 ± 1.0σ. Sound given alone 16 times during training period, C–R's occurred 14 times, median time 122 ± 2.2σ. For unconditioning, sound given alone 23 times, responses occurred in all except 1st, 3d, 17th, and 18th trials, median time 163.0 ± 31.5σ. Finally, 25 measures made of V. R. T., median time 225.0 ± 11.7σ. *S* said the very loud sounder frightened him at first and that the resulting start jerked his finger off the key. The V. R. T. measured at the beginning of the experiment (135 ± 2σ) was quite fast and regular, and the V. R. T. measured at the end of the experiment (225 ± 12σ) was slow and less regular. One peculiarity of this experiment is that C–R time (122σ) was faster than time of response to sound and shock (129σ). After the experiment had been completed, *S* said his hand was withdrawn reflexly in the C–R's and that he did not notice it until it was already off the key. He said, however, that he was reacting voluntarily during the unconditioning. The sound and shock seemed to come at the same time during the training period.

Exp. 10, Mr. T. A.—Twenty-five measures made of V. R. T. to loud telegraph sounder, median time 294.0 ± 12.2σ. Training stimuli: loud sounder and shock, sounder given 20σ before shock, stimuli repeated 215 times, median time of response to both stimuli 142.5 ± 1.1σ. Sound given alone 25 times during training period, C–R's occurred in last 22 trials, median time 257.0 ± 25.2σ. For unconditioning, sound given alone 14 times, responses occurred on all trials, median time 245.0 ± 12.3σ. Finally, 25 measures made of V. R. T., median time 216.0 ± 8.8σ. After the experiment had been completed, *S* said he had been reacting voluntarily on all occasions when the sound stimulus was given alone.

Exp. 11, Miss E. S.—Thirty measures made of V. R. T. to electric bell, median time 195.0 ± 6.6σ. First training stimuli: bell and shock, bell given 20σ before shock, stimuli repeated 109 times, median time of response to bell and shock 170.0 ± 2.6σ. Bell given alone 11 times during training period, C–R occurred only on 2nd trial, time of this C–R 222σ. Second training stimuli: telegraph sounder and shock, sounder given 20σ before shock, stimuli repeated 141 times, median time 143.0 ± 9.3σ. Sounder given alone nine times during training period, C–R occurred only on 3rd trial, time of this C–R 131σ. Third training stimuli: sounder and shock given simultaneously 36 times, median time of response to sounder and shock 153.5σ. Shock given alone 17 times during this training period, median time 205σ; and sound given alone two times during this same training period, and no C–R on either trial. Finally, 30 measures made of V. R. T. to sounder, median time 169.5 ± 1.1σ. *S* said she just felt no tendency to react in most of the C–R trials.

Exp. 12, Mr. G. B.—Thirty measures made of V. R. T. to telegraph sounder, median time 145.5 ± 4.7σ. Training stimuli: sounder and shock, sounder given 40σ before shock, stimuli repeated 216 times, median time of response to both stimuli 141.0 ± 1.6σ. Sounder given alone 24 times during training period, C–R's occurred in all trials, median time 135.0 ± 2.8σ. For unconditioning, sounder then given alone 17 times, responses occurred in all trials, median time 162.0 ± 5.1σ. Finally, 25 measures made of V. R. T. to sounder, median time 141.0 ± 6.0σ. One peculiarity of this experiment is that time of C–R's to sound (135σ) was faster than time of responses to sound and shock (141σ). After the experiment had been completed *S* said that in the successful C–R trials he got his finger off the key before he knew it, but that during unconditioning he was reacting voluntarily.

In Table 1 we have summarized all experiments except the sixth, and the table is largely self-explanatory. This summary brings out clearly several irregularities and inconsistencies in the results. C–R time is not always faster than V. R. T., and in several experiments the difference between

TABLE 1

GENERAL SUMMARY OF RESULTS

(Sound Stimulus and Shock were the Training Stimuli in Each Experiment. Time is Median Time in σ.)

Exp. No.	Subject	Sound Stimulus	V. R. T.	Time of Giving Sound in Training	No. Repetitions of Training Stimuli	Time of Responses to Sound and Shock	Time of Responses to Shock Alone	% of C–R's to Sound Alone	Time of C–R's	% of Responses in Unconditioning	Time of Responses in Unconditioning	V. R. T.
1	Mr. J. H.	Sounder		With Shock	323	180±1		92	183±3	58	189±12	252±28
2	Mr. B. C.	Sounder	192±6	With Shock	335	166±1	168±3	100	169±6			199±8
3	Miss E. B.	Relay		With Shock	163	174±2		88	188±13	75	194±6	274±17
4	Miss H. S.	Relay	306±6	With Shock	220	232±2		92	245±18	53	299±28	279±9
5	Mr. L. G.	Relay	185±4	With Shock	323	143±1		97	157±2	31		215±5
7	Mr. J. W.	Relay	219±6	With Shock	135	145±1		13				199±4
8	Mr. A. R.	Sounder	169±5	10σ Before Shock	223	145±2		100	188±12			225±7
9	Mr. L. A.	Sounder	135±2	20σ Before Shock	160	129±1		88	122±2	83	163±32	225±12
10	Mr. T. A.	Sounder	294±12	20σ Before Shock	215	143±1		88	257±25	100	245±12	216±9
11	Miss E.S.	Bell Sounder Sounder	195±7	20σ Before Shock 20σ Before Shock With Shock	109 141 36	170±3 143±9 154	 205	9 11 0				 170±1
12	Mr. G. B.	Sounder	146±5	40σ Before Shock	216	141±2		100	135±3	100	162±5	141±6

C–R time and the time of the responses in unconditioning is small. The time of the responses to sound and shock is on the whole faster than the time of C–R's to sound alone, but in two experiments the time of C–R's to sound alone is faster than the time of responses to sound and shock.

Interpretation of the Results

From what is known about the factors involved in the present experiment and from the verbal reports of our *S*'s, it appears that soon after they were placed in the experimental situation they formed a verbal and motor set or disposition to withdraw their hand when they heard the sound. This reaction tendency remained voluntary in the majority of our experiments, but the sometimes involuntary nature of the hand withdrawal movement has been frequently described in the literature on reaction time (**14**, pp. 177–202; **6**; **2**; **12**).

The factor of expecting the shock increased the speed of hand withdrawal movements, and this factor was present in every trial when the speed of conditioned hand withdrawal movements was measured. Some of the C–R's which we obtained were voluntary reactions, and others were at least in some measure involuntary. Hamel (**8**) compares the fastest human reflexes with conditioned hand withdrawal movements, and he also makes the unjustified assumption that all reflexes are fast and non-conscious. Our C–R times were considerably faster than those obtained by Hamel. A fast time may eliminate the possibility of a voluntary response, but a slow time does not eliminate the possibility of an involuntary response. Hunter and Hudgins (**9**) have recently given an interesting treatment of the nature of voluntary activity.

Verbal activities seem to play a part in motor learning that is not always recognized (**5**), and the changes which from time to time take place in the *S*'s verbal activities can readily account for the irregularities in our results on conditioned hand withdrawal responses. On the whole the most efficient way to get a person to withdraw his hand when he hears a sound is to give him the verbal instructions to withdraw his hand when

he hears the sound. In the conditioned hand withdrawal experiment with a person as the subject, the person's verbal activities will be present, and they will be active in a different way at different times, regardless of whether or not he is told that he is supposed to learn something and regardless of whether or not he is told what it is that he is supposed to learn. It would seem that more emphasis should be placed on the role of verbal activities in motor learning than is customary in most of the discussions of this subject.

(Manuscript received May 7, 1934)

REFERENCES

1. BECHTEREW, W., *La psychologie objective*, tr., 1913.
2. BREITWIESER, J. V., Attention and movement in reaction time, *Arch. of Psychol.*, 1911, No. 18.
3. CASON, HULSEY, The conditioned eyelid reaction, J. EXPER. PSYCHOL., 1922, 5, 153–196.
4. CASON, HULSEY, A note on the conditioned eyelid reaction, J. EXPER. PSYCHOL., 1923, 6, 82–83.
5. CASON, HULSEY, The role of verbal activities in the conditioning of human subjects, *Psychol. Rev.*, 1934, 41, 563–571.
6. FROEBERG, SVEN, The relation between the magnitude of stimulus and the time of reaction, *Arch. of Psychol.*, 1907, No. 8.
7. GINSBERG, LEON, Difference discrimination and the conditioned reaction, Columbia University Ph.D. thesis, 1924
8. Hamel, I. A., A study and analysis of the conditioned reflex, *Psychol. Monog.*, 1919, 27, No. 118.
9. HUNTER, W. S., AND HUDGINS, C. V., Voluntary activity from the standpoint of behaviorism, *J. Gen. Psychol.*, 1934, 10, 198–204.
10. RAZRAN, G. H. S., Conditioned responses in children: a behavioral and quantitative critical review of experimental studies, *Arch. of Psychol.*, 1933, No. 148.
11. RAZRAN, G. H. S., Conditioned withdrawal responses with shock as the conditioning stimulus in adult human subjects, *Psychol. Bull.*, 1934, 31, 111–143.
12. TODD, J. W., Reaction to multiple stimuli, *Arch. of Psychol.*, 1912, No. 25.
13. WATSON, J. B., The place of the conditioned-reflex in psychology, *Psychol. Rev.*, 1916, 23, 89–116.
14. WOODWORTH, R. S., AND POFFENBERGER, A. T., Textbook of experimental psychology (mimeographed), 1920.

READING AND LEGIBILITY OF CHINESE CHARACTERS: IV. AN ANALYSIS OF JUDGMENTS OF POSITIONS OF CHINESE CHARACTERS BY AMERICAN SUBJECTS

BY SIEGEN K. CHOU[1]

National Tsing Hua University,
Peiping, China

Introduction.—In the third paper of this series on Reading and Legibility of Chinese Characters, two experiments had been reported on judging positions of Chinese characters by American subjects.[2] *Experiment I* was an intensive individual experiment with the author's Quadrant Bradyscope[3] through simultaneous, voluntarily maintained exposure[4] of 64 Chinese characters by and to 7 subjects themselves. This first experiment revealed the fact that while correct judgments of the four possible positions of Chinese characters, *i.e. upright, upside-down, turned-to-the-right,* and *turned-to-the-left,* amounted to about 65 percent, the incorrect judgments ranged from about 26 percent for *Opposite* errors, about 5 percent for *Analogous* errors, and about 4 percent for *Crossed* errors.[5] This inequality of prevalence of the different kinds of incorrect judgments needed further analysis as to its cause. Why is there a tendency on the part of foreigners to judge

[1] The writer is indebted to Mr. Chao-Tang Lei for the preliminary statistical calculations and tabulations, upon which all the final percentages reported in this paper were recalculated by the writer himself.

[2] This paper is the fourth of a series on Reading and Legibility of Chinese characters. The first three appeared in this JOURNAL: I. Influence of reading-direction and character-position upon speed, 1929, **12**, 156–177; II. Reading half-characters, 1930, **13**, 332–351; III. Judging the positions of Chinese characters by American subjects, 1930, **13**, 438–452.

[3] S. K. Chou, A quadrant tachistoscope for studying the legibility of Chinese characters, this JOURNAL, 1929, **12**, 178–186.

[4] S. K. Chou, Reaction-keys and a new technique for reading-reactions, *Amer. J. Psychol.*, 1929, **41**, 469–473. Cf. Tachistoscope versus Bradyscope, *ibid.*, 1930, **42**, 303–306.

[5] *Op. cit.*, III, pp. 443–446. For the definition of these errors, see p. 443.

characters in various positions more often in the geometrically *Opposite* direction or position than in the logically *Analogous* or *Crossed* position?

It is for this purpose that *Experiment II*, an extensive group test, was also conducted and reported, which consisted of a sheet of 312 Chinese characters printed in various positions to be marked out by 142 American subjects with no knowledge at all of the Chinese language.[6] The mean group score of correct judgments is 149 characters with a probable error of 3 characters, which is about 48 percent accuracy. Preliminary analysis of written, subjective comments from the subjects brought out a number of interesting objective cues that were actually used by them for determining the true position of a Chinese character.[7] These may be listed according to their degree of importance as follows:

(1) All horizontal lines have a little bump on the upper part of the stroke.
(2) All vertical lines are much heavier and all horizontal ones much lighter.
(3) The base of the character is broad, wide, open, straight, flat, even, long, heavy, dark, solid, light, small, and separated.
(4) The general appearance may be one-sided, natural, typical Chinese, emphasis, logical, being upright, best, looking right, feel, judgments, square, and *Gestalt*.
(5) The top of the character is short, light, fine, sharp, square, pointed, broad, heavy, wide, open, and flat.
(6) The manner of writing is also a criterion: tapered stroke, stroke direction, shaded at top of stroke, shaded at bottom of stroke, and curved down and out.
(7) Resemblance to or association with objects, English, and the very letter 'T' used for marking the true position of a Chinese character.
(8) Guess or chance, balance or stability.
(9) Symmetry or proportion.
(10) Knowledge of the language.

[6] *Op. cit.*, III, pp. 446–452. For the Group Character-Position Test Sheet, see the original reproduction Fig. III, p. 447.

[7] *Op. cit.*, III, pp. 449–451.

Problem.—The object of this paper is to analyze further some of the objective determining factors inherent in the characters themselves as suggested by the above subjective comments. Specifically we want to verify the foreigners' subjective impressions of the characters, to determine the intrinsic characteristics of Chinese characters that condition correct judgment of their true position or uprightness; to see especially (1) whether the ratio of the number of horizontal and vertical strokes has any appreciable effect on correct judgment, (2) the alleged effect of the characteristic horizontal stroke upon judgment, and (3) whether general form, pattern or configuration conditions accuracy of judgment; and finally to compare the results of this less rigidly controlled group test with that of the first relatively more rigidly controlled individual experiment.

Statistical Procedure.—All judgment markings, on the test sheets, of the various positions of every one of the 312 Chinese characters by 142 subjects were carefully counted and tabulated by several clerks [8] in the following manner:

(1) One clean sheet of the original printed copy was first taken for marking out all the 78 upright characters, which were then all cut out so that 78 little windows appear in their places. This is then the marking or scoring key for all upright characters.

(2) Another three printed copies were similarly cut for upside-down, turned-to-the-right, and turned-to-the-left characters so that one scoring key may be prepared for each of the positions. (It was discovered later that the simplest character for 'one,' which is merely a horizontal dash, was wrongly printed turned-to-the-left, when it ought to be, as a matter of fact, turned-to-the-right. So there were actually only 77 characters turned-to-the-right but 79 characters turned-to-the-left. This, however, has practically no effect upon our subsequent calculations in view of the fact that the number of cases are rather large.)

(3) Then the scoring key sheet for upright characters, say, was placed upon the test copies of different subjects one by one and the actual judgments of every upright character by each subject were read off and copied down in blanks. The symbols used for recording are the respective initial or suggestive letters of the words or expressions denoting character-position, *i.e.*, '*u*,' '*d*,' '*r*,' '*l*.' Two persons performed this task, one calling out the symbols and the other recording them in blanks.

(4) All judgments of the other three positions of characters were similarly recorded down.

(5) Then the number of '*u*,' '*d*,' '*r*' and '*l*' judgments were counted separately for each character for every actual position of the characters.

---

[8] Thanks are due to Mr. Lung Chukao of the *Research Committee in Education and Psychology*, The National Association of the Mass Education Movement, Ting Hsien, China, for the assistance of several clerical helps.

(6) Out of the 142 subjects who passed judgments on the test, 104 or 73.3 percent marked every one of the 312 characters, while 38 or 26.7 percent only marked a varying portion of the total number. These were tabulated separately and labelled respectively Group I and Group II in the following treatment of results, while Group III is the result of the entire group of 142 subjects.

The above procedure is necessarily tedious and error-inducing. But since the relative proportions of the number of judgments are quite variable in size for the different judgments of various actual positions of the characters, our use of the relative percentage between judgments in the following treatment seems to justify our negligence of the possible and probable errors involved in the large numbers of repeated counting and recording. These mechanical errors, if present (and they must be), are certainly less than one tenth of a percent.

The distribution of the various number of judgments for the various actual character-positions is shown in Table 1.

TABLE 1

NUMBER OF JUDGMENTS BY 142 AMERICAN SUBJECTS OF 312 UPRIGHT, UPSIDE-DOWN, RIGHT-TURNED, AND LEFT-TURNED CHINESE CHARACTERS

Group	Nature of Judgments	Character-Positions			
		Upright	Top-Down	R-turned	L-turned
I (Complete) $N=104$	*Upright*	*4636*	2563	1497	1898
	Top-Down	1871	*3894*	1121	746
	R-turned	904	724	*3464*	1907
	L-turned	692	907	1922	*3639*
	?—doubtful	9	25	4	22
	Total No.	8112	8113	8008	8212
II (Incomplete) $N=38$	*Upright*	*1331*	727	453	596
	Top-Down	648	*1260*	441	283
	R-turned	390	366	*1087*	588
	L-turned	311	328	670	*1240*
	? and No Ans.	284	282	275	299
	Total No.	2964	2963	2926	3006
III (Whole) $N=142$	*Upright*	*5967*	3290	1950	2494
	Top-Down	2519	*5154*	1562	1029
	R-turned	1294	1090	*4551*	2495
	L-turned	1003	1235	2592	*4879*
	? and No Ans.	293	307	279	321
	Total No.	11076	11076	10934	11218

The completely marked test copies were separately treated from the incompletely marked ones, and the entire of these were then grouped together. This is labelled in the Table as Group I, II, & III respectively. The effect of this separation of treatment for the completely and incompletely marked records lowers the percentages of certain judgments but increases those of certain others; but the general tendency of the relative percentages between various judgments are strikingly similar, showing the fact that reliability is increased as the number of characters checked is increased. Table 2 shows the corresponding percentages of Table 1.

TABLE 2

PERCENTAGE OF JUDGMENTS BY 142 AMERICAN SUBJECTS OF 312 UPRIGHT, UPSIDE-DOWN, RIGHT-TURNED, AND LEFT-TURNED CHINESE CHARACTERS

Group	Nature of Judgments	Character-Positions			
		Upright	Top-down	R-turned	L-turned
I (Complete) $N=104$	*Upright*	*57.1*	31.6	18.7	23.1
	Top-Down	23.3	*48.0*	14.0	9.1
	R-turned	11.1	8.9	*43.3*	23.2
	L-turned	8.4	11.2	24.0	*44.3*
	?—doubtful	.1	.3	.0	.3
	Total %	100.0	100.0	100.0	100.0
II (Incomplete) $N=38$	*Upright*	*44.8*	24.5	15.5	19.9
	Top-Down	21.9	*42.5*	15.1	9.4
	R-turned	13.2	12.4	*37.1*	19.6
	L-turned	10.5	11.1	22.9	*41.2*
	? and No Ans.	9.6	9.5	9.4	9.9
	Total %	100.0	100.0	100.0	100.0
III (Whole) $N=142$	*Upright*	*53.8*	29.7	17.8	22.2
	Top-Down	22.7	*46.6*	14.3	9.2
	R-turned	11.7	9.8	*41.7*	22.2
	L-turned	9.1	11.2	23.7	*43.5*
	? and No Ans.	2.7	2.7	2.5	2.9
	Total %	100.0	100.0	100.0	100.0

Frequency of Judgments.—The relative frequencies of the various judgments vary both as to actual character-positions and according to completeness of record. Very few doubtful judgments were encountered in the complete records, which are mostly due to the fact that the symbol 'T' was written in

such a confused fashion under some characters that its true position could not be legibly determined, and hence the positions of the characters denoted themselves also could not be determined. But in the incomplete records, unanswered or doubtfully answered characters amounted to 9.6, 9.5, 9.4, and 9.9 percents for *Upright*, *Top-Down*, *R-turned*, and *L-turned* characters respectively. However, when the complete and incomplete records were pooled together, the percentages reduce all to about 3 percent.

Upright characters elicited 57.1 percent *Upright* judgments for those who completely checked the words, but a little less, namely, 44.8 percent, for those who partially checked most of the characters. The pooled result for the entire group of 142 subjects gives 53.8 percent judgments identical with the true character-position. The reverse *Top-Down* judgments are the next most frequent, while *R-turned* and *L-turned* judgments follow in frequency. This is true for both the complete, incomplete, and whole groups I, II and III.

However, this relative frequency of judgments is radically changed as the true character-position is changed. Thus, we see that the most frequent judgment for *Upside-down* characters is *Top-Down* judgment, for *R-turned* characters *R-turned* judgments, and for *L-turned* characters *L-turned* judgments. The next most frequent judgment is the reverse of the most frequent in every case. This result seems to indicate that there must be some naturalness about the uprightness of a Chinese character to foreigners and also that this natural uprightness may be interpreted wrongly in the reverse. Just why this is so must be further elucidated from the qualities of the characters themselves.

Moreover, there is also some consistency in the two least frequent judgments for every true character-position. Judging from these relative frequencies, *R-turned* and *L-turned* judgments are perhaps both unnatural judgments for *Upright* characters; but this unnaturalness seems to be less in the former than in the latter kind, because the former percentage is consistently higher than the latter. This is also true with *Upside-down* characters, but the relative percentage is re-

versed in size except in the Incomplete Group. In other words, whereas for *Upright* characters *R-turned* judgments seem to be more natural, for *Top-Down* characters *L-turned* judgments seem to be more so, because the percentages are in each case higher, no matter whether we base our comparison on Complete data or on the Whole Group including some Incomplete data.

This uniformity of result is even more striking in the case of *R-turned* and *L-turned* characters. Here it is the *Upright* and *Top-Down* judgments that seem to be unnatural, but the unnaturalness is much less for *Upright* than for *Top-Down* judgments, especially when the actual position of the characters is *L-turned*, because the percentages are much larger for both Complete, Incomplete, and Whole Groups of data (23.1, 19.9, and 22.2 percent respectively).

The above results show us that positional judgments by foreigners of the uprightness of Chinese characters purposely distorted are by no means random guesses. If these Americans were somehow or other not sincere in their judgments but marked the test copies blindly simply because they were asked to do the psychologist a favor, the percentages should come out more or less uniform in size, *i.e.* 25 percent. In fact, the purely chance judgments should be exceptionally close to the probable 25 percent, because the number of judgments *in toto* for every actual character-position is more than eleven thousand for the Whole Group, about three thousand in the Incomplete Group, and no less than eight thousand in the Complete Group. That the 142 American subjects taken as a whole were rather sincere in passing their judgments upon every one or most of the 312 Chinese characters purposely distorted as to their uprightness is not only conceivable but statistically valid. Now the question is how did they perform this seemingly obvious and easy task but really marvellous and interesting feat, especially when we remember that none of these Americans had any previous knowledge at all of the Chinese language.

Nature of Judgments.—Before we go to the very construction and architecture of Chinese characters themselves for

explanations or cues of judgments, we must first analyze further the different classes of judgments and see what general tendencies are revealed perhaps less obviously in them.

Results from Individual Bradyscopic Test through Simultaneous Voluntary Maintenance of Exposure in Experiment I had already furnished data for a classification of all judgments into four kinds [9] as follows:

(1) Correct or identical judgments
  (*a*) Upright as Upright
  (*b*) Upside-Down as Top-Down
  (*c*) R-turned as R-turned
  (*d*) L-turned as L-turned
(2) Incorrect or dissimilar judgments
  (*a*) Opposite judgments
    1. Upright as Top-Down and *vice versa*
    2. R-turned as L-turned and *vice versa*
  (*b*) Analogous or homolateral judgments
    1. Upright as R-turned and *vice versa*
    2. Upside-Down as L-turned and *vice versa*
  (*c*) Crossed or diagonal judgments
    1. Upright as L-turned and *vice versa*
    2. Upside-Down as R-turned and *vice versa*

A comparison of the relative percentages of the various classes of judgments both for the Individual Experiment and for the Group Test may be seen in Table 3, where the Complete and Incomplete data are also separated. The relative percentages for the various classes of judgments are a little higher for Complete data (Group I) than for Incomplete data (Group II), which fact is explicable by the general tendency of increased reliability due to increased number of cases (here positional judgments per individual). The pooled percentages from all data are all lower than those of Group I but higher than those of Group II. In general, results from Individual Experiment agree strikingly well with those from Group Test, *Correct* judgments being the most frequent, *Opposite* judgments next frequent, and *Analogous* and *Crossed* judgments the least frequent.

[9] *Op. cit.*, III, p. 443.

TABLE 3

COMPARISON OF PERCENTAGE OF VARIOUS CLASSES OF POSITIONAL JUDGMENTS OF CHINESE CHARACTERS BY GROUP AND INDIVIDUAL TESTS ON AMERICAN SUBJECTS

Group	Correct		Opposite		Analogous		Crossed	
	Judgments	%	Judgments	%	Judgments	%	Judgments	%
I ($N=104$)	U—u	57.1	U—d	23.3	U—r	11.1	U—l	8.4
	D—d	48.0	D—u	31.6	D—l	11.2	D—r	8.9
	R—r	43.0	R—l	24.0	R—u	18.7	R—d	14.0
	L—l	44.3	L—r	23.2	L—d	9.1	L—u	23.1
	Mean	48.1	Mean	25.5	Mean	13.6	Mean	12.5
II ($N=38$)	U—u	44.8	U—d	21.9	U—r	13.2	U—l	10.5
	D—d	42.5	D—u	24.5	D—l	11.1	D—r	12.4
	R—r	37.1	R—l	22.9	R—u	15.5	R—d	15.1
	L—l	41.2	L—r	19.6	L—d	9.4	L—u	19.9
	Mean	41.4	Mean	22.2	Mean	12.3	Mean	14.5
III ($N=142$)	U—u	53.8	U—d	22.7	U—r	11.7	U—l	9.1
	D—d	46.6	D—u	29.7	D—l	11.2	D—r	9.8
	R—r	41.7	R—l	23.7	R—u	17.8	R—d	14.3
	L—l	43.5	L—r	22.2	L—d	9.2	L—u	22.2
	Mean	46.4	Mean	24.6	Mean	12.5	Mean	13.9
Individual Experiment	U—u	75.0	U—d	21.4	U—r	1.8	U—l	1.8
	D—d	60.7	D—u	26.8	D—l	5.4	D—r	7.1
	R—r	57.2	R—l	29.4	R—u	9.8	R—d	3.6
	L—l	66.6	L—r	25.4	L—d	3.6	L—u	4.5
	Mean	64.8	Mean	25.9	Mean	5.1	Mean	4.2

Note.—The capitals in the table denote actual Character-Positions: U = Upright; D = Upside-Down; R = Right turned; L = Left turned. The corresponding small letters: u, d, r, and l denote the corresponding judgments. Group I are those 104 Ss who completely marked the 312 Chinese characters; Group II are those 38 Ss who partially or incompletely marked the characters; while Group III are the entire 142 Americans who answered the Group Test.

Looking closer into the Mean percentages in Table 3, we note the fact that, in the Incomplete data Group II, the relative size of the percentages of the four classes of judgments has the tendency to gravitate toward the most probable number. Although this tendency is still present in the Whole data Group III and even in the Complete data Group I, it is perhaps significantly altered by Individual Experiment. Thus, we see that *Correct* judgments increased from 41.4 percent for Incomplete data to 48.1 percent for Complete data, al-

though the latter is somewhat lowered when the pooled Whole group yielded an average of 46.4 percent. It is interesting to note that percentage of *Correct* judgments increased to 64.8 in the Individual Experiment. This is perhaps due to the fact that conditions were more or less better controlled in the Individual Experiment than in the Group Test. That this higher percentage of *Correct* judgments were not obtained at the expense of *Opposite* judgments is indicated by the fact that the percentage of the latter remains practically the same as in the Group Test.

The percentage of *Correct* judgments is increased in the Individual Experiment rather at the expense of *Analogous* and *Crossed* judgments than at the expense of *Opposite* judgments. This is shown by the large decrease of the mean percentages from Group Test to Individual Experiment in *Analogous* and *Crossed* judgments. There are 12.5 percent *Analogous* judgments in the entire group of 142 test records, but in the Individual Experiment, the percentage is only 5.1; and the decrease from 13.9 percent to 4.2 percent in the case of *Crossed* judgments is even more noticeable.

It may be inferred from the above results that the accuracy of positional judgments of Chinese characters by American subjects is not only a function of the degree of control of experimental conditions but also of the nature of judgments as well. *Correct* judgments tend to increase when individual laboratory experiment is used rather than group or individual paper-pencil test. *Opposite* judgments have the same tendency although to a lesser extent. The increase in frequency of both *Correct* and *Opposite* judgments is curiously all at the expense of *Analogous* and *Crossed* judgments, *i.e.* both the latter classes of judgments decrease as the former increase, when method of investigation is changed from paper-pencil test to individual laboratory experiment.

It should be noted that whereas *Correct* judgments are judgments identical with the true position of the character judged, *i.e.* without distortion in spatial order, *Opposite* judgments are, however, judgments equivalent to that position of the character where it were as if rotated 180 degrees and

Analogous and *Crossed* judgments are judgments with only 90 degrees distortion. Now just why is it that Americans (or foreigners in general?) tend to pass more judgments rotated 180 degrees from the true position than those rotated only 90 degrees from it? It is perhaps interesting to know what effect intermediate positions between 90 and 180 degrees would have upon judgment frequency and accuracy. This is of course another problem for which our present data can not furnish an answer. But it is clear that fallible judgments of purposely tilted Chinese characters in various positions (here only the four cardinal positions were used) somehow depend upon the tilted position itself and that disagreement between judgment and true position is more frequent in the case of 180 degrees than in the case of 90 degrees, correct judgments being, of course, of 0 degrees distortion.

Character-Gestalt and Accuracy of Judgment.—The writer had pointed out in an earlier discussion [10] the significance of *word-* or *character-Gestalt in a line* in effecting the speed of reading in Chinese subjects. Here, in the present case, individual Chinese characters were purposely printed in various positions so that the regular, natural *line-Gestalt* of all upright characters is distorted. The task is, however, only a discrete appraisement of the individual characters for *uprightness*. As we have shown above, this *uprightness* judgment varies with initial true position, method of measuring, and nature of judgment. The degree of rotation of the characters is especially significant for accuracy and frequency of judgments. These extrinsic factors are perhaps only secondary in importance in comparison with the intrinsic qualities of the Chinese characters themselves, which we may conveniently term the characteristic *character-Gestalt.*

In order to test this hypothetical influence of *word-Gestalt* upon accuracy of judgment, we first select ten characters in each position that give the highest accuracy and another ten in each position that give the lowest accuracy. We then compare the two series of characters and find out whether there

[10] S. K. Chou, *Gestalt* in reading Chinese characters, *Psychol. Rev.*, 1930, **37**, 54–70. Cf. *op. cit.*, I, pp. 168–172.

TABLE 4

PERCENTAGES OF CORRECT JUDGMENTS FOR 10 MOST ACCURATELY JUDGED AND 10 LEAST ACCURATELY JUDGED CHARACTERS IN EACH TRUE POSITION

CHARACTERS		CHARACTER POSITIONS							
		Upright		Top-Down		R-turned		L-turned	
		Word	%u	Word	%d	Word	%r	Word	%l
MOST ACCU-RATELY JUDGED	1	禽	90.7	齊	75.1	午	72.1	會	80.8
	2	西	76.0	久	69.8	乃	67.8	用	73.6
	3	爪	77.1	炊	70.2	再	64.8	兒	65.2
	4	賽	75.0	人	69.1	布	64.8	穌	64.8
	5	十	76.0	行	67.9	復	65.4	花	63.1
	6	鬪	72.6	食	68.4	屋	65.8	下	62.4
	7	不	71.3	赫	67.2	條	60.0	香	62.4
	8	芙	72.5	谷	66.3	街	56.9	示	63.5
	9	白	72.3	宿	65.6	卿	57.4	殲	59.4
	10	人	72.0	內	65.6	国	58.2	實	57.6
	Mean %		75.6		68.8		63.2		65.3
LEAST ACCU-RATELY JUDGED	11	逃	40.0	喜	30.3	鬱	25.5	樹	32.4
	12	瞭	40.0	垂	29.6	鬱	25.3	矛	31.5
	13	誨	40.0	多	28.9	鑄	24.5	憐	30.5
	14	阜	39.3	鄧	27.4	鈎	23.8	殴	30.0
	15	針	39.3	十	25.9	亞	24.2	世	29.9
	16	逮	37.0	干	25.2	紛	24.2	卸	29.9
	17	誌	36.1	咽	22.2	頗	21.9	立	27.9
	18	比	33.3	目	18.0	再	20.0	主	23.0
	19	輕	28.1	科	17.0	紙	18.9	才	19.4
	20	到	21.4	臣	16.3	鷗	17.4	一	15.5
	Mean %		31.8		24.3		22.6		26.0

are characteristic differences in the general shape of the individual characters. The percentages of correct judgments for ten most accurately judged and ten least accurately judged characters in each true position are shown in Table 4.[11]

The mean percentages of the 10 most accurately judged characters for each true position are all more than twice as large as those of the 10 least accurately judged characters. The 10 *Upright* characters which average 75.6 percent correct judgments are comparatively simpler in structure, more symmetrical (without a single exception), and fewer in the number of strokes than the 10 least accurately judged characters which average only 31.8 percent. The 10 least accurately judged *R-turned* characters which average only 22.6 percent correct judgments are especially complicated in structure, numerous in the number of strokes, and uniform in the fact that all except the fifteenth character are vertically separable into two right and left halves that are not symmetrical. On the other hand, the percentages of correct judgments seem to be correlated negatively with the number of strokes in the 10 most accurately judged *R-turned* characters but positively in the 10 least accurately *L-turned*. Thus we see that the least accurately judged Chinese character (only 15.5 percent) is the twentieth, a simple horizontal stroke standing for 'ONE' that is *L-turned*, the one above it in the same position having only three strokes; but the two most accurately judged characters (72.1 and 67.8) respectively) in the *R-turned* position have the least number of strokes. Stability seems to be the most important characteristic for judging the *upside-downness* of Chinese characters, because the 10 most accurately judged *Top-Down* characters which average 68.8 percent correct judgments all look like standing on heads; while the 10 least accurately judged ones do not suggest this feeling of unstability at all. In fact, the 10 *Top-Down* characters that average only 24.3 percent accuracy are mostly equally wide from top to bottom of the character and hence very hard to tell whether upside-

[11] For the original test sheet of Chinese characters in various orientations, see *op. cit.*, III, Fig. III, p. 447.

down or not, especially the last four characters which score only 22.2, 18.0, 17.0, and 16.3 percents respectively.

TABLE 5

COMPARISON OF PERCENTAGES OF CORRECT JUDGMENTS OF CHARACTER-POSITIONS IN TWO EXPERIMENTS UNDER DIFFERENT CONDITIONS

Various Conditions	Character-Positions			
	Upright	Top-Down	R-turned	L-turned
40 characters (among 312) most accurately marked by 142 Ss as to true position	75.6	68.8	63.2	65.3
		(Mean = 68.2)		
64 characters correctly judged in the Bradyscope by 7 Ss	75.0	60.7	57.2	66.6
		(Mean = 64.8)		
312 characters correctly marked by 142 Ss as to true position	53.8	46.6	41.7	43.5
		(Mean = 46.4)		
40 characters (among 312) least accurately marked by 142 Ss as to true position	31.8	24.3	22.6	26.9
		(Mean = 26.4)		

Referring to Tables 3, 4, and 5, we note the fact that the 40 most accurately judged characters in various positions average 68.2 percent correct judgments, even higher than in the case of much more controlled individual bradyscopic experiment, where only 64.8 percent accuracy was secured from 7 subjects judging 64 characters in various positions. Moreover, the percentage of accuracy of judgment for the 40 best characters is much higher than that for all the 312 characters pooled together (only 46.4 percent), although the 40 worst characters average much lower, namely 26.4 percent. It is therefore evident that Chinese characters present enormous individual differences in the degree of difficulty for judging *uprightness* and this difficulty is conditioned by such factors as method of judging, manner of control, etc.

It is the usual contention of both Chinese and foreign scholars that the Chinese language, being non-alphabetical, is the hardest language to learn, because of the fact that each

character has to be learned, memorized, and recalled individually. Sweeping generalizations as to this imagined difficulty of the language being the cause of high percentage of illiteracy in China are more superficial than real. The writer had already pointed out in the first paper of this series that "earlier discoveries by Cattell and Dodge of the importance of the total word-unit, by Goldscheider and Müller, and Zeitler of certain determining and dominating parts in every word, and by Messmer and Huey of the dependence of the total character of a word upon the width, height and geometrical form of the letters, and of the varying importance of the first and last, upper and lower halves of a word, should point in the right direction for making a really pertinent analysis of the corresponding effects and differences in the case of the Chinese language." [12] It is fallacious to think that the Western languages are easier to learn (even this is highly hypothetical) simply because they are alphabetic or syllabical. Experiments have repeatedly demonstrated to us that people do not perceive individual letters or syllables as such but a whole word or several words as a unit. This fact has resulted in a complete revolution in our methods of teaching children to read. Children now learn from the start to recognize whole words organized in meaningful sentences.[13] Are we Chinese justified then to abandon the Chinese characters altogether in favor of some alphabetic language?

It seems to the present writer that the very essence of the Chinese character lies in the individualistic, typical, and characteristic form of each character, which is a much better, more precise, intimate, and closed configurational structure than any Western alphabetical word composed of a number of letters-in-line. The relative ease for reading in the two cases, *i.e.* Chinese characters *vs.* English word, had already been demonstrated previously in a theoretical discussion entitled '*Gestalt in Reading Chinese Characters.*' [14] The writer wishes only to point out the fact that, if improved methods can be

[12] S. K. Chou, *Gestalt . . . , op. cit.*, pp. 156–157.

[13] A. I. Gates, Psychology and education, in *Psychology at Work*, edited by P. S. Achilles, 1932, p. 45.

[14] *Op. cit.*

devised for teaching Chinese characters, the very nature of this language is, perhaps, not only conductive to learning, memory, and recall in itself but possibly more so than any Western alphabetic language. It should be pointed out that the Chinese language is fundamentally a visual language rather than an auditory one.[15] Being such, its evolution through different stages of pictorial representation has brought the individual characters to such a degree of perfection that each one of them is a good, closed, and coherent visual *Gestalt* for perceptual reading. That this is evident may be demonstrated by even presenting Chinese characters to foreigners merely to be judged of the uprightness of each character, where numerous *Gestalt*-qualities inherent in the single character itself determine in the main accuracy of judgment. Four of the more important *Gestalt*-qualities of Chinese characters may be briefly mentioned.

(1) *Symmetry.*—The two most accurately judged characters have, among other good qualities, a perfect symmetry. One of them being *Upright* (Table 4, No. 1) is judged correctly by 90.7 percent of the 142 Americans, and the other being *L-turned* has also 80.8 percent accuracy. The 10 *Upright* characters ranging from 90.7 to 72.0 percent accuracy are all symmetrical without a single exception. At least 7 out of the 10 most accurately judged *Upside-down* characters are symmetrical, namely, No. 1, 2, 4, 6, 7, 8, and 10, although only two or three in the other two positions are symmetrical. On the other hand, if we look over the lower portion of Table 4, where the 40 least accurately judged characters are represented, majority of them are unsymmetrical except *Upside-down* No. 11, 12, 16 and 18, *R-turned* No. 15, and *L-turned* No. 17, 18, and 20. The importance of this factor will be further indicated presently.

(2) *Organization.*—Symmetry is only one aspect of good organization. A character may be perfectly symmetrical but lacks organization and hence is hard to judge as to its true

[15] Shih Hu, in an autobiographical sketch of the history of Chinese Literary Renaissance, *Eastern Miscellany* (Chinese), 1934, Vol. 31, No. 1, pp. 15–31, seems to maintain that the very visual character of the Chinese language constitutes the fundamental objection.

position. For example, the *Upright* No. 14, 18, the *Upside-down* No. 15, 16, 18, and 20, the *L-turned* No. 17, 18, 19, and 20 are all more or less symmetrical but lack somewhat unity and coherence. This is especially striking in the *L-turned* character No. 19 and 20 and also in the *Upside-down* characters No. 15 and 16. These characters are too simple, entirely devoid of any trace of organization. Therefore they are comparatively less accurately judged as to their true position.

(3) *Proportion.*—By proportion we mean the even distribution of strokes in a character. Of course all symmetrical characters are more or less proportional. But there are characters that cannot be strictly classified as symmetrical. Nevertheless the strokes are evenly distributed on the top or bottom, right or left of the character, betraying no sign of over-balance on any one side. Thus, we see that the *Upright* characters No. 19 and 20 are over-balanced on the left side and No. 12 and 13 on the right, the *Upside-down* No. 14 and 17 on the right, the *R-turned* No. 20 and the *L-turned* No. 14, 16 on the left. Hence they are comparatively less accurately judged. On the other hand, the first two *R-turned* characters, although neither symmetrical nor full in organization, score 72.1 and 67.8 percent accuracy respectively, because the few strokes are evenly and proportionally distributed.

(4) *Stability.*—All characters that are broad, open, flat, straight, heavy, dark, solid, and even at the bottom tend to be more stable. Most of the good characters mentioned above are also in good equilibrium, *e.g.* the *Upright* No. 1 and the *L-turned* No. 1. The character for 'MAN' (*Upright* No. 10) is especially wide at the base. Several of the *Upside-down* characters (*e.g.* No. 2, 3, 4, and 8) are decidedly wider at the base. These are all in the upper half of Table 4, the most accurately judged characters. Casual observation of the lower part of the table is enough to convince us that none of the 40 least accurately judged characters possess this stability quality. A sense of equilibrium is felt also when the character is pointed like the pyramid at the top and supported by two parallel vertical strokes at the base (No. 1 *Upright* and *L-turned*). In such characters as *Upright* No. 4, *Upside-*

down No. 1, and *L-turned* No. 10, although the top seems to be overloaded, the two left and right pointing strokes at the bottom, suggesting the human feet on ground, compensate this effect more or less.

The two groups of good and bad characters shown in Table 4 may be explained by one or more of the above four *Gestalt*-qualities. Characters that are most accurately judged by foreign observers tend to be symmetrical, organized, proportional, and stable; those that are least accurately judged seldom possess these qualities, except when symmetry is present to such a degree (*e.g. Upside-down* characters No. 16 and 18, and *L-turned* No. 17, 18, and 20) that it becomes a hint for incorrect opposite judgments.

The effect of these *Gestalt* qualities upon foreign judgments of the mere *Uprightness* of Chinese characters may be still more evident when we analyse the nature of the wrong judgments for the worst characters. The average percentage of each of the three types of wrong judgments, namely, *Opposite*, *Analogous*, and *Crossed*, for each of the four positions, when the worst five characters in each case are considered, is between 26 and 56 percent, while pooled result from all characters gives an average number of wrong judgments of all sorts between 9 and 30 percent only (Table 3, Group III). This strong tendency in some characters toward certain definite type of positional error may be explained by the above *Gestalt*-qualities. Reference should be made to Table 6 in connection with this discussion. Since the *Opposite* wrong judgments of the position of 60 most easily interpreted characters tend to increase to 43, 56, 42, and 46 percent for *Upright*, *Upside-down*, *R-turned*, and *L-turned* characters respectively, which are all nearly twice as frequent as in the pooled result of all 312 characters as a whole, there must be some inherent quality in these ambiguous characters that make foreigners hard to decide whether or not they should be turned 180 degrees around. Symmetry seems to be the important factor that conditions high percentage of wrong *Opposite* judgments. This is explained by the fact that, when a character is symmetrical, it is usually identical in its

TABLE 6

COMPARISON OF PERCENTAGE OF VARIOUS CLASSES OF WRONG JUDGMENTS OF THE POSITIONS OF CHINESE CHARACTERS

Characters	Opposite		Analogous		Crossed	
	Judgments	%	Judgments	%	Judgments	%
60 characters most frequently judged wrong by 142 Ss as to true position	U—d	43	U—r	29	U—l	26
	D—u	56	D—l	28	D—r	26
	R—l	42	R—u	36	R—d	31
	L—r	46	L—d	28	L—u	43
312 characters as a whole judged wrong by 142 Ss.....	U—d	23	U—r	12	U—l	9
	D—u	30	D—l	11	D—r	10
	R—l	24	R—u	18	R—d	14
	L—r	22	L—d	9	L—u	22

right and left halves and also tends to be equally wide from top to bottom: hence it is easy to pass oppositely wrong judgments. Other characters among the 60 give *Analogous* and *Crossed* judgments more or less near the chance 25 percent except in the case of 'L—u' *crossed* judgment, which is as high as 43 percent. This last exception is due to the fact that two characters (*L-turned* No. 19 and 20, Table 4) are most frequently judged *Upright* by the 142 Ss (87 and 89 persons respectively).

The above detailed analysis revealed the fact that both correct and wrong judgments of the true positions of Chinese characters may be attributed to some inherent *Gestalt*-qualities of the characters themselves, among which *symmetry*, *organization*, *proportion*, and *stability* have been mentioned. All these qualities had been commented upon by the American subjects who marked the characters. Thus we see that subjective impressions of foreigners on the general make-up of Chinese characters are objectively borne out in the actual records. The inference is, then, that general form, pattern, or organization of the Chinese character conditions accuracy and error of position judgment. This may be further shown although less clearly by an actual classification of the 312 characters according to certain general principles of structure in the characters themselves.

Character-Structure and Accuracy of Judgment.—The architecture of the Chinese character follows certain principles of structural make-up. Majority of the characters used in this test is *symmetrical, i.e.* the right and left halves are exactly identical, if an imaginary line is drawn vertically through the middle of each one, no matter whether the character is actually possible of bisection or not. Some examples are given already above (see Table 4). In those that are separable into parts, both horizontal and vertical division is possible. A large number of characters are *vertically separable* into two or three parallel parts or units, which are either of approximately the same number of strokes or of unequal weight either in terms of total lineal length of strokes or in terms of total area occupied in the left, middle, or right portion of the character. Examples are *Upright* characters No. 12, 13, 15, 17–20; *Top-Down* characters No. 5, 7, 17, 19; *R-turned* characters No. 1, 5, 7–9, 11–14, 16–20; and *L-turned* No. 4, 9, 11, 13, 14, 16. *Horizontally separable* characters are also very many, such as *Upright* No. 4, 8, 14; *Top-Down* No. 1, 9, 11; and *L-turned* No. 5, 7, 8, 10, and 18. There are other characters, however, that are *indivisible* by any of the above criteria but may assume a perfect, *square-shaped* whole, or give the suggestion of *left-bottom supporting* right-top, although some may be *diagonally divisible* (either left-top to right-bottom or *vice versa*). These classifications may be set in a table as follows.

TABLE 7

A Classification of Chinese Characters According to Structure

Name of Classification	Examples	Words
1. Diagonally Divisible	*Top-Down* No. 3, 13	18
2. Symmetrical	*Upright* No. 1, 5, 10	71
3. Horizontally Separable	*L-turned* No. 7, 8, 10	67
4. Indivisible	*L-turned* No. 6, 19	46
5. Vertically Separable	*R-turned* No. 1, 5, 11	118
6. L-bottom supporting	*Upright* No. 16	12
7. Square-Shaped	*Top-down* No. 10, 18	7

The 312 Chinese characters may be classified into the above seven classes. But 27 of the *symmetrical* ones also belong to several of the other classes and hence the total number of words in Table 7 is 339, not 312. The examples all refer to the characters in Table 4.

It has been pointed out above that *symmetry* is the most important *Gestalt*-quality that conditions correct judgments of character position. This quality together with *organization*, *proportion*, and *stability* refers, however, to the totality of a character. We have reason to suspect that the part-whole relationships as listed in Table 7 may also condition accuracy of judgment. So we regrouped the thousands of judgments according to the above classification and also according to the nature of judgment and calculated the different percentages as in the following Table 8.

TABLE 8

COMPARISON OF PERCENTAGE OF VARIOUS KINDS OF POSITIONAL JUDGMENTS OF CHINESE CHARACTERS CLASSIFIED ACCORDING TO NATURE OF PART-WHOLE RELATIONSHIPS

Classification	Correct	Opposite	Analogous	Crossed
Diagonally Divisible	56.3	17.0	13.2	13.5
Symmetrical	52.2	27.8	8.4	11.5
Horizontally Separable	50.7	26.6	10.3	12.4
Indivisible	45.9	28.9	15.6	9.6
Vertically Separable	43.5	23.2	16.4	16.9
L-bottom Supporting	43.2	33.1	11.4	12.3
Square-Shaped	36.4	34.3	12.4	16.9
Mean	46.9	27.3	12.5	13.3

Character-structure is perhaps a conditioning factor of accuracy of judgment. Characters *diagonally divisible* score 56.3 percent *Correct* judgments; those that are *symmetrical* and *horizontally separable* have all above 50 percent accuracy; the next three classes of characters, namely, *indivisible*, *vertically separable*, and *left-bottom supporting*, all above 43 percent; although only the *square-shaped* ones score as low as 36.4 percent. This influence of character-structure upon judgment may be further evidenced by the fact that the relative percentages of *Opposite* judgments are just the reverse of correct judgments. In other words, when one kind of character gives the highest percentage of correct judgments of its true position, the opposite kind of judgment is the smallest in percentage for that character. Thus we see that *diagonally divisible* characters elicit 56.3 percent *Correct* judgments with

only 17.0 percent incorrect *Opposite* judgments, while *square-shaped* characters, yielding only 36.4 percent *Correct* judgments, induced 34.3 percent incorrect *Opposite* judgments. For the other kinds of characters, the same relationship holds for all except *symmetrical* and *vertically separable* characters, where the percentages of wrong *Opposite* judgments are 27.8 and 23.2 respectively, the former being too high and the latter too low.

It is natural that *squared-shaped* characters ought to give the least cue for correct judgment of their true position, because the characters are mostly ambiguous as to uprightness. The top and bottom are so much alike that one can hardly tell the head and tail of a character and hence it is most often judged in the opposite position. Characters that are more or less an *indivisible* whole are harder to tell their true position than those *diagonally divisible*, *symmetrical*, and *horizontally separable*. But why *vertically separable* characters are even still harder to tell cannot be explained off-hand.

It will be recalled that American subjects reported two important cues upon which they mostly based their positional judgment of Chinese characters. These are (1) that *all horizontal lines have a little bump on the upper part of the stroke* and (2) that *all vertical lines are much heavier and all horizontal ones much lighter*. If these important cues subjectively noted as criteria of judgment are really effective, the relative percentage of the various classes of judgment must bear this out. It is perhaps reasonable to suspect that percentage of correct judgment ought to increase as the number of horizontal lines, and hence the number of the characteristic little bumped arrow-heads, increases. There are 67 characters whose number of horizontal strokes is from zero to six in access of vertical strokes. Table 9 shows the relation between difference in horizontal and vertical strokes and percentage of various judgments.

Judging from the percentage of *Correct* judgments, accuracy seems to decrease, contrary to what is expected, as the number of horizontal-minus-vertical strokes increases, although the result is by no means consistent. Thus we see

TABLE 9

COMPARISON OF PERCENTAGE OF VARIOUS KINDS OF POSITIONAL JUDGMENTS OF CHINESE CHARACTERS CLASSIFIED ACCORDING TO DIFFERENCE BETWEEN HORIZONTAL AND VERTICAL STROKES

Difference between Horizontal and Vertical Strokes	Correct	Opposite	Analogous	Crossed
19 characters 0 dif.	45.7	24.1	14.2	16.0
18 characters 1 dif.	47.7	23.9	14.5	13.9
18 characters 2 dif.	41.2	34.2	12.2	12.4
18 characters 3–4 dif.	47.3	26.2	13.1	13.4
10 characters 5–6 dif.	36.6	24.7	18.3	20.4
Mean	42.9	26.6	14.5	15.2

that 18 characters having one more horizontal than vertical strokes yield 47.7 percent *Correct* judgment, but 10 characters differing in 5–6 strokes have only 36.6 percent accuracy. This general tendency is true for all characters except *Upside-down* ones, where accuracy seems to increase as the number of horizontal-minus-vertical strokes increases. The percentages of correct judgments for *Upside-down* characters having a horizontal-minus-vertical difference of 0–1, 2, 3–4, and 5–6 strokes are 34.9, 37.6, 39.7, and 43.8 respectively.[16] Since the number of characters differing in horizontal and vertical strokes is only from 10 to 19, the above result is perhaps not significant at all. The same thing may be said of the relation between the number of little bumped, arrow-headed horizontal strokes and accuracy of judgment, the different percentages of which are shown in Table 10.

Judging from the relative percentage of *Correct* judgments, perhaps no significant difference should be attached to the number of bumped, arrow-headed horizontal strokes in a Chinese character, when foreigners are asked to guess the distorted true position of it. All the percentage differences from one to five bumps are better interpreted as chance differences, although the special low percentage of *Correct* judgments for six-to-eight bumps, namely, 35.8 percent, may be explained by the fact that the number of characters is too small in comparison with others. The number of characters having

[16] Not shown in the table.

one, two, three, four, five, and six-to-eight bumped, arrow-headed horizontal strokes as listed in Table 10 is 81, 69, 34, 25, 14, and 10 respectively.

It must be pointed out that, although both the character-

TABLE 10

COMPARISON OF PERCENTAGE OF VARIOUS KINDS OF POSITIONAL JUDGMENTS OF CHINESE CHARACTERS HAVING DIFFERENT NUMBERS OF BUMPED, ARROW-HEADED HORIZONTAL STROKES

Number of Bumped, Arrow-headed Horizontal Strokes	Correct	Opposite	Analogous	Crossed
One Bumps	51.3	24.1	10.9	13.7
Two Bumps	45.3	25.9	14.1	14.7
Three Bumps	46.5	23.8	15.2	14.5
Four Bumps	40.3	25.7	17.8	16.2
Five Bumps	47.6	21.9	13.3	17.2
Six–Eight Bumps	35.8	26.4	19.4	18.4
Mean	44.5	24.6	15.1	15.8

istic bumped, arrow-headed horizontal stroke and the ratio between horizontal and vertical strokes in a Chinese character had been noted most frequently by American subjects to be two important cues in guessing the uprightness of the characters, the relation between these two factors and uprightness was assumed by some subjects wrongly. As noted in the earlier report, no fewer than a dozen subjects worked on an entirely wrong assumption. The one who, marking all 312 characters, made a score of 31 had only one cue and that was just the reverse of correct. The most important correct cue is that *all horizontal lines have a little bump on the upper end of the stroke.* He and two others took this to be at the lower left end and hence the absurdity. This most important cue has been taken also as turned to the right, *i.e.* it is upright when the line is vertical with the bump on the top. About 8 made this assumption. About 18 who discovered this as a cue took it in the right position. The other obvious cue is that *all vertical lines are much heavier and all horizontal ones much lighter.* About 24 sensed this but there was one who took just the opposite hypothesis.[17] It is of little wonder, then, that the objective

[17] *Op. cit.*, III, p. 450.

record does not reveal any beneficial effect of these two characteristics of the strokes of Chinese characters upon positional judgment.

The Psychology of Chinese Characters.—To recapitulate, foreigners judging the true position of abnormally orientated Chinese characters do not give more correct judgments when they consciously note certain characteristic elements of the character than when they are not aware of anything particular at all. Structural make-up of the character in the sense of *separability* into smaller units either *diagonally*, *horizontally*, or *vertically*, for example, has doubtful influence on accuracy of positional judgment. Even such consciously noted characteristics as the little bumped, arrow-headed horizontal stroke and the relative frequency of horizontal and vertical strokes show no appreciable effect on accuracy at all. On the other hand, such vague intrinsic qualities of the characters themselves as *symmetry*, *organization*, *proportion*, *stability*, and *squareness* all tend to show great effect on accuracy of judgment. This is perhaps some indication that the psychology of Chinese characters is one example of the Psychology of *Gestalt*.

The psychology of Chinese characters has three aspects, namely, legibility, perception, and reading. Treated just as nonsense symbols, Chinese characters are good material for experimental purposes. Reading involves percetion of individual characters but both reading and perception are conditioned by legibility. I have endeavored in several places [18] to direct the attention of investigators to the legibility aspect of the problem. Such factors as form or individual differences of the characters, size of the character, style of the character, position of the character in the group, spacing between characters, length of line, leading, appearance of Arabic digits or foreign words, punctuation and special marks, and idex system are important taken by themselves. Perceptual studies of Chinese characters on foreign subjects may contribute much toward this problem of legibility of Chinese characters.

[18] *Op. cit.*, I, pp. 156–158.

The earliest experimental research on the psychology of Chinese characters was done by Lew between 1916 and 1919 [19] on 9 American adults and 140 American children, beside 98 Chinese adults and 18 Chinese children. It deals with learning, reproduction, recall, and recognition. Among the conclusions reached by Lew from his results, the following have some bearing on our problem. Lew concluded that the arrangement of strokes is one of the important factors in the process of learning the characters; that a change in the arrangement of strokes has a decided influence over the mastery of the characters so far as writing them is concerned; that the number of units and the arrangements of the units have a determining influence over the process of learning; that the number of strokes a character has does not tell the degree of difficulty it will present to the learner; that the number of strokes in a character has a far greater influence over the form of a character in learning it than over the sound or meaning; and that the number of strokes of a character does not increase or decrease the degree of difficulty of a character—not because the number of strokes does not have any influence, but rather because of the greater influence of some other factors, among which are the similarity of forms, the usefulness of homonyms, the occasional identity between sound and meaning, the suggestiveness of the forms, and the representation of meaning by the common simple radicals.[20]

Evidently ignorant of Lew's study, Tsai and Abernethy performed an experiment on 'Complexity of Characters and Difficulty of Learning,' [21] which consists of presenting 30 characters with arbitrary English equivalents 3 seconds each to 21 American subjects for recognition, and 5 seconds each to 9 others for reproduction. Recognition took place in 4 seconds and reproduction in 10 both after three presentations and after twenty-four hours. The thirty characters are in

[19] T. T. Lew, The psychology of learning Chinese characters, *Chinese Soc. & Pol. Sci. Rev.*, 1923, 7, No. 4, 120–181; 1924, 8, No. 1, 1–53; No. 2, 1–80; No. 3, 114–219; No. 4, 1–105. Also separate monograph, pp. 377.

[20] *Ibid.*, 1924, 8, No. 4, 100–105; Monograph, pp. 374–377.

[21] L. S. Tsai and E. Abernethy, The psychology of Chinese characters, this JOURNAL, 1928, 11, 430–442.

three series of 10 each consisting of a varying number of strokes of 3, 6, and 12 respectively. The number of strokes is taken as an indication of complexity against the finding of Lew's results. Although difficulty in learning was found out not related to complexity, a logarithmic equation was derived by the method of least squares for the function between reproduction (by writing) and complexity (three, six, and twelve strokes only.[22]

Hull used twelve series of twelve Chinese characters each in an experiment on the evolution of concepts.[23] The 144 characters shown in Plate I of Hull's monograph are mostly not true Chinese characters, but "freely adapted" symbols resembling such. The symbols within any one series have one common radical but the common radical is increasingly obscured from the first to the twelfth symbol. The twelve different radicals serve as concepts, the evolution of which was the prime purpose of the investigation. The characters as shown in Hull's monograph are terribly distorted and mutilated out of all proportion, symmetry, and balance. It must be noted that Chinese characters in their proper proportion and balance are seldom copied correctly even by the Chinese student.

Kuo used 88 compound characters, of which 75 belong to 5 common radicals, each radical having 15 characters.[24] Just as in Hull's experiment, the response consisted of generalizing the common structure of the compound characters—namely, the radical—and of the common meaning or common relationship which this radical represents. The remaining 13 characters are negative instance characters outside the radical groups. The components of the 88 compound characters together with 24 new characters were written on another sheet of card-board for testing the recognition of the radicals and pseudo-radicals, for testing inference, and for testing the detec-

[22] *Ibid.*, p. 435–436.

[23] E. L. Hull, Quantitative aspects of the evolution of concepts, *Psychol. Monog.*, 1920, 28, No. 123, pp. 86. The *adapted* symbols resembling Chinese characters are reproduced on p. 10.

[24] Z. Y. Kuo, A behavioristic experiment on inductive inference, this JOURNAL, 1923, 6, 247–293. Two samples of Chinese characters are reproduced in Plate I, p. 251, and Plate II, p. 254.

tion of negative instance. There is more to be learned about the recognition and reproduction of these numerous characters in both Hull's and Kuo's reports than in Tsai and Abernethy, whose experiment was devised especially for that purpose.[25]

These researches utilizing Chinese characters as material of experiment all bring out either directly or indirectly the fact that the perception of a Chinese character is concordant with the laws of *Gestalt.* Chinese as well as foreigners respond not to the specific elements of the individual character as such but to the total form, configuration, or *Gestalt* of the character. No matter whether the subject is asked to learn, to perceive, to recognize, to reproduce, or simply to tell the abnormally orientated position, as is reported in this and the earlier reports of the present writer, he is always guided more by the total qualities of the individual character rather than by the particular, peculiar, characteristic element within it. The individual symbolic word itself is unitary and, when standing side by side with a row or column of several others, forms still greater unitary line-whole or *Gestalt*, as already pointed out elsewhere.[26] It is only natural, then, that the psychology of Chinese characters is, in reality, a Psychology of *Gestalt* in particular, into which we cannot go any further in this connection.

Summary and Conclusions.—The data of an extensive group test on 142 American subjects on the problem of guessing or judging purposely abnormally orientated Chinese characters in the upside-down, right-turned, and left-turned, as well as upright, positions were further analyzed in some detail. Preliminary results reported earlier show that the mean group score of correct judgments of 312 abnormally orientated Chinese characters is 149 characters, which is about 48 percent accuracy. Most of the subjectively commented cues used in guessing are borne out by the objective records, although a few specifically noted characteristics of the constituent strokes do not seem to be conditioning factors on accuracy of judgment.

[25] Cf. M. Wylie, Recognition of Chinese symbols, *Amer. J. Psychol.*, 1926, 37, 224–232.

[26] S. K. Chou, *Gestalt . . . , op. cit.*, pp. 156–157.

(1) Both the characteristic little bumped, arrow-headed horizontal stroke of a Chinese character and the ratio between the number of the characteristically different horizontal and vertical strokes have no appreciable effect on correct judgment of uprightness of Chinese characters. If there is any effect, it is detrimental to judgment accuracy, because the subjects have the chance of taking a wrong assumption concerning these two most frequently noted cues.

(2) Several *Gestalt*-qualities of the individual characters may be taken to account for the high percentage of correct judgment of the uprightness of Chinese characters, among which are symmetry, organization, proportion, and stability or balance. The influence of these factors clearly shows that the psychology of Chinese characters is but one instance of the general Psychology of *Gestalt*. Subjective comments of the individual characters as *unitary wholes*, such as that the base of the character is broad, wide, open, straight, flat, even, long, heavy, dark, solid, light, small, and separated; that the general appearance may be one-sided, natural, typical Chinese, emphasis, logical, being upright, best, looking right, feel, judgments, square, and the very word *Gestalt;* that the top of the character is short, light, fine, sharp, square, pointed, broad, heavy, wide, open, and flat; that the manner of writing is also a criterion: tapered stroke, stroke direction, shaded at top of stroke, shaded at bottom of stroke, and curved down and out; and that resemblance to or association with objects, English, and the very letter 'T' used for marking the true position of the character—such subjective comments of foreigners cannot, of course, be ascribed to the constituent parts of a Chinese character.

(2) On the other hand, such arbitrary part-whole classifications of the 312 Chinese characters into such structural make-up as separability into smaller units either diagonally, horizontally, or vertically, and inseparability, squareness, or left-bottom of a character supporting right-top do not give us any significant difference in the accuracy of positional judgment of abnormally orientated Chinese characters.

(3) That the total *Gestalt*-qualities of a Chinese character

determine its perception, recognition, reading, and legibility is supported incidentally by several related studies such as Hull's experimental study of the quantitative aspects of the evolution of concepts, Kuo's behavioristic experiment on inductive inference, Lew's learning of Chinese characters, Wylie's study on recognition of Chinese symbols, and Tsai and Abernethy's article on the psychology of Chinese characters. Since Chinese characters frequently appear in experimental reports, foreign investigators who wish to utilize such material should well bear the above facts in mind.

THE BIOCHEMICAL VARIABILITY OF THE INDIVIDUAL IN RELATION TO PERSONALITY AND INTELLIGENCE [1,2]

BY HYMAN GOLDSTEIN

Columbia University

It is becoming increasingly apparent that a great many psychological phenomena must be interpreted from physiological viewpoints. The organism is to be regarded as an aggregate of physico-chemical systems functioning in dynamic relationship to each other. The field of biological chemistry looms large as the possible final court wherein many of the perplexing problems of personality and abnormal psychology will ultimately be settled.

The necessity for a biochemical approach to the study of personality has been well voiced by Rich (**25**). Any attempt that can be made to reduce the aberrations of personality into physiological terms which may in turn be stated in the quantitative language of biochemistry must be gratifying and requires no apology.

The problem of this study is to determine whether personality and intelligence have biochemical correlates. More concretely stated, it is the purpose of this investigation to see whether the personality trait scores determined by the Bernreuter Personality Inventory and the test scores [3] determined by the Thorndike CAVD Intelligence Test, Levels M-Q, bear any relationship to certain constituents of the blood.

[1] The writer desires to acknowledge the debt that he owes to Dr. O. Klineberg for advice and constructive criticism throughout the investigation; to Prof. R. S. Woodworth and to Prof. A. T. Poffenberger for many helpful suggestions; to Prof. H. T. Clarke and to Prof. E. G. Miller for the laboratory facilities and for invaluable suggestions offered to him at the College of Physicians and Surgeons.

[2] This investigation is based upon the data reported by: Goldstein, H. A biochemical study of the metabolism of mental work. *Arch. Psychology*, 1934, No. 164.

[3] The test scores reported for the CAVD test are the summated scores and not the level scores. The results of this study are the same regardless of the form in which the CAVD results are expressed.

1. Personality

The biochemical literature relating to investigations of personality has for the most part been confined to studies of neurotic and psychotic individuals. While the list of studies cited below makes no claim to completeness, it does indicate how conflicting the results obtained have been.

Many studies have been made of the hydrogen ion concentration of the body fluids of neurotic and psychotic individuals. Some of the investigators have shown that the excited type of individual has, in general, body fluids of low hydrogen ion concentration while the lethargic type has, in general, body fluids of high hydrogen ion concentration.

Ludlum (**18**) has presented some clinical evidence for two types of acute insanity: one, the excited type with alkaline saliva and alternatingly acid and alkaline urine, perspiration and feces; the other, the confused type with a general acid predisposition. Starr (**30**) differentiated a lethargic type of stammerer with a distinctly acid saliva from an excitable type with a neutral or alkaline saliva, that is, there was a negative correlation between acidity and excitability in the individual. Obraszov (**20**) studied the alkaline reserve of a group of psycho-neurotic children and concluded that these children possessed a high alkaline reserve.

An investigation of particular interest to psychologists is reported by Rich (**25**) who made a study of the urine and saliva samples of fifty-seven University students. Each subject gave three samples, each one week apart. The subjects were rated by fraternity brothers on the traits of good-naturedness, perseverence, leadership, aggressiveness and emotional excitability. There was also another group of 303 delinquent children from whom 1 blood sample was obtained. These children were rated by four persons (psychological examiner, physician, psychiatrist, and laboratory technician), with whom they came in contact during the routine examinations, on traits of good-naturedness, aggressiveness, and emotional excitability. In the first group the hydrogen ion concentration of the saliva and the acidity of the urine samples were determined. The values obtained from the three sam-

ples were averaged in each case. The alkali reserves of the blood samples were also determined. Rich concludes that the least excitable individuals tend to have the most acid urine while more excitable persons tend toward neutrality or alkalinity of their fluids; that the least aggressive subjects tend to excrete the greatest amounts of acid in urine and blood while more aggressive individuals show opposite characteristics; that the degree of relationship found to exist between chemical determinations and personality traits is expressed by correlation coefficients between + .20 and + .30.

Suitsu (**31**) found a greater variability in plasma CO_2 combining capacity in a group of psychotics than in a group of normal subjects.

Other investigators have failed, however, to find any indication of relation between hydrogen ion concentration of body fluids and neurosis or psychosis.

Suitsu (**31**) found that the alkaline reserve (as measured by the amount of CO_2 in the blood plasma combined as bicarbonate) of the insane falls within the normal limits for healthy persons. Excited patients did not differ from depressed patients in the amount of alkalinity. Henry and Mangam (**15**) in a study of 200 patients consecutively admitted to a hospital have found that the CO_2 combining power is unaffected by psychosis unless a definite physical disease accompanies the psychosis.

Ballif and Gherscovici (**2**) studied 300 cases of nervous and mental diseases and found that neither the hydrogen ion concentration nor the alkali reserve showed significant variations from the normal acid-base balance. Appel, Farr, and Hodes (**1**) reached the conclusion that the total base content of the blood plasma and the plasma CO_2 were normal in manic-depressive psychoses.

Gaskill (**12**) found that the hydrogen ion concentration of successive saliva samples varied significantly under rest conditions.

Investigations have also been made of other blood determinations than hydrogen ion concentration of the body fluids of neurotic and psychotic individuals. On the one hand,

some of the studies have shown that the fluids of these individuals deviate from the normal in various constituents.

Weinberg (**34**) found an increased excretion of creatinine in urine from patients suffering from nervous disorders as compared with a group of normal patients. The patients had been placed on a diet free from creatine and creatinine. Pighini (**23**) investigated the cholesterol values of serum in different forms of insanity and concluded that in manic-depressive insanity and in dementia præcox high values prevailed. Looney and Childs (**17**) found that schizophrenic patients tended to have a lower blood cholesterol content than a group of normal controls. Schube (**29**) in a study of 200 neuropsychiatric patients and 54 normal individuals found that the former group had a significantly lower blood cholesterol content than the latter. The patients also had a greater variability than the normal individuals. Cannon (**7**, p. 66) has shown that sugar metabolism plays a part in emotional disturbance.

Rich (**25**), in the study mentioned above, also determined the inorganic phosphorus values of the urine samples of the 57 university students. The values obtained from the three samples given by each student were averaged in each case. The inorganic phosphorus and creatinine of the blood samples of the 303 delinquent children were determined. Rich found that the emotionally excitable individuals have less creatinine in blood and urine; that phosphorus metabolism appears to be bound up with personality in some way, the details of which are not clear; that the degree of the relationships found to exist between chemical determinations and personality traits is expressed by correlation coefficients between + .20 and + .30. Rich (**26**) in a study of the 171 twenty-four hour urine specimens furnished by the 57 students found significant correlations between the total acid and creatinine, between the creatinine and phosphorus, and between the total acid and phosphorus.

Hammett (**14**) made a study of the relative variability of the soluble nitrogenous compounds more commonly associated with protein metabolism together with the total nitro-

gen and sugar of the blood. He had 12 subjects who were hospital patients and nurses. None of the subjects suffered from metabolic disorders. Three of the subjects gave 6 blood samples each while the other nine subjects gave 4 blood samples each. The blood from any single individual was not taken more often than once in seven days. Hammett (**14**) concludes that while the total nitrogen, non-protein constituents, and the sugar of the blood vary in the same individual from week to week, there is a tendency for the level of these variations to be characteristically individual, that is, individuals might be differentiated by the amount of variation.

Hammett (**13**), studied the relation between emotional stability and metabolic stability in 17 individuals among whom were included emotionally stable and unstable insane and normal individuals. Classification of the emotional status of the subjects was arrived at by the consultation with the physicians in actual contact with the patients and by personal observation over considerable periods of time. Analyses were made at weekly intervals for total nitrogen, non-protein nitrogen, urea nitrogen, creatinine, creatine, uric acid, amino acids, and sugar. The coefficient of variability for each blood constituent for each individual was calculated. The sum of these coefficients was taken as the total variability of the intermediary metabolism of the individual. There was found a relation between a relatively high metabolic stability and a high emotional stability and between low metabolic stability and emotional excitability. As mentioned above, Suitsu (**31**) found that the variability of the plasma CO_2 combining capacity is higher in the insane than in normal individuals.

In this case also, other investigators have shown that neurotic and psychotic individuals do not differ from normal individuals in the concentrations of these constituents of the body fluids.

Bloor (**6**) in a study of 100 epileptic and 32 normal subjects found no significant difference between the blood plasma cholesterol values of the epileptics and those of the normal individuals. Weston (**35**) reported that the blood of epileptic,

dementia praecox and manic-depressive patients showed no deviation from the normal content of creatinine, creatine, glucose, chlorides or calcium. Barrett and Serre (**3**) analyzed the blood of 20 manic-depressive and 20 dementia præcox cases and found normal figures for creatinine and glucose. Henry and Mangam (**15**) found normal blood glucose and chloride values in a group of psychotic individuals. Pighini (**23**) found that in senile dementia præcox and in most cases of epilepsy the serum cholesterol values were normal. Looney and Childs (**17**) found no relationship between blood cholesterol content and ratings of emotional status in a group of schizophrenic patients. They concluded that no single determination of cholesterol can be taken as characteristic since the variability is so great and the range is so wide in normal and in abnormal persons. Simple determinations of the cholesterol content of the blood were found to be worthless as chemical aids except in extreme conditions.

Patrick and Rowles (**21**) found no relationship between the scores of 52 university women on the Clark revision of the Thurstone Personality Schedule and their basal metabolic rates.

2. Intelligence

The biochemical literature relating to differences in intelligence levels, while less voluminous, is no less conflicting in its results.

Some investigators have found relationships between intelligence differences and biochemical differences.

Powers (**24**) made blood tests on 12 normal or superior persons and 20 idiots. All were adults of approximately the same age and health and each group included an equal number of males and females. The calcium of the blood serum and the inorganic phosphorus of the blood plasma were determined. Every case, normals and idiots, came within the normal range in calcium content. Without a single exception the idiots had a high phosphate content, averaging 8.95 mg per 100 cc blood as compared with an average of 4.36 mg for the normal group. Rich (**25**) found a correlation of $-.51$ between intelligence and excretion of phosphorus per kilo of body weight in urine in 39 subjects.

Other investigators, however, fail to find any connection between intelligence and the biochemical values of the body fluids.

Negus (**19**) took blood samples from 25 idiots, aged 11–44 years. He failed to find abnormal quantities of inorganic phosphate in their blood, the average being 4 mg per 100 cc of blood. While the acid-soluble phosphates, hemoglobin, red cell count and cell volume were normal, there was indication of low cholesterol values in the blood of idiots. Pighini (**23**) found that the cholesterol and oxycholesterol values of blood serum were normal in imbecility.

Rich (**27**) made studies of the blood of a random sampling of children referred to the Illinois Institute for Juvenile Research. The sampling was limited by the success of the technician in securing sufficiently large blood specimens from each individual. The children included intelligence levels ranging from the imbecile to the superior. Rich determined the blood calcium, inorganic phosphorus, total phosphorus and lipoid phosphorus and found no correlation between intelligence and phosphorus in any form or between intelligence and calcium content.

Patrick and Rowles (**21**) found no relationship between the scores of 52 university women on the Ohio State University Intelligence Test and their basal metabolic rates.

Procedure: Part I

The experimental work is divided into two parts. The subjects of Part I were 19 adult males of college rank, ranging in age from 18 years 6 months to 32 years 11 months, averaging 23 years 9 months.

It was deemed advisable to secure blood samples instead of urine specimens for biochemical analysis. Rowe (**28**, p. 146) has stated that: "The blood is the great thoroughfare of the body both for the ingress of absorbed materials and the egress of body wastes. The delicately balanced regulating mechanisms of the body maintain the concentrations of certain important constituents at surprisingly constant levels when they are determined under definite conventional limitations. To this fact is attributable the very real significance that attaches to such levels."

The determinations made included inorganic phosphorus, sugar, cholesterol, creatinine, calcium and chlorides. The inorganic phosphorus content of the blood plasma was determined by the Youngburg (**37**) method. The sugar content of the whole blood was determined by the Folin-Wu (**10**) method. Bloor's (**5**) method was used in measuring the cholesterol content of the whole blood. The Folin-Wu (**9**) method was used to determine the creatinine content of the whole blood. The Clark-

Collip (8) modification of the Kramer-Tisdall (16) method for calcium and the Wilson-Ball (36) modification of the Van Slyke (33) serum chloride method were employed. Duplicate determinations were not performed in the experimental analyses. However, preliminary experimentation showed that quadruplicate samples of the same specimen checked highly.

The Bernreuter Personality Inventory (4) which measures several different aspects of personality at one time was used. The traits measured are neurotic tendency, self-sufficiency, and dominance-submission. The Personality Inventory may be scored separately on each of these 3 scales. The scales possess high reliability. No coefficient of reliability is below + .85 and the average is + .88. The inventory possesses a high degree of validity, that is, the three traits measured by the Personality Inventory are identical with the three traits which have been measured by previously validated tests. No corrected coefficient of validity is below + .84 and the average is + .97. Since in the construction of the Personality Inventory the previously validated tests, such as, the Thurstone Neurotic Inventory, the Bernreuter Self-Sufficiency Test, and the Allport Ascendance-Submission Reaction Study, were used to locate individuals who possessed the various traits to an extreme degree and since the weights on the scoring keys of the Personality Inventory were computed on the basis of the extent to which each question differentiated between the criterion groups composed of these extreme individuals (4), it is evident that the coefficients of validity are 'stacked' and hence spuriously high.

The Thorndike CAVD Intelligence Scale, Levels M–Q (32) furnishes a more extensively graduated range of difficulty than any other intelligence scale. It is superior to the usual intelligence examination since individuals may be compared with respect to absolute ability in the functions measured by the scale (11). The levels M–Q are particularly adapted to college students and only about 20 percent of college graduates pass level Q (get more than 50 percent correct). The reliability is about + .80 and the validity is about + .70 (32, p. 99).

The subjects reported at the laboratory without breakfast at 9 A.M. Sunday morning. They waited at least 20 minutes in a chair. An 18 cc blood sample was then taken by a physician. Each subject came to the laboratory three times at intervals of three weeks.

It is obvious that the subjects were not under laboratory controlled situations during the intervals separating the experimental sessions and hence the experimenter knew neither the kind nor the amount of activity that each subject had participated in between the times that he reported at the laboratory. In order to see what the effect of a laboratory controlled situation would have upon the blood values in contrast to a noncontrolled situation, it was thought advisable to determine the effects of the passage of an interval of three hours during a state of rest and very little overt muscular activity. For this purpose it was necessary to have a control group.

Nine subjects out of the group of nineteen served as a control group in a period of relaxation. In order to increase the reliability of the results, it was deemed necessary to increase the size of the control group. However, since the remainder of the original group of nineteen subjects was unavailable the control group of nine subjects was increased to fourteen subjects by five additional male subjects who were, in general of the same age range, educational, social, and economic status as that of the original group of nineteen subjects. The control group reported without breakfast and reclined in Morris chairs and on sofas. The subjects of this group were urged to sleep if possible, or if that were impossible, to read light literature, magazines or newspapers. This was

continued for three hours. Blood samples were taken before and immediately after the session.

Results

As stated above, the subjects came directly to the laboratory from their homes without breakfast and waited at least twenty minutes in chairs. While it is evident that these may not be ideal basal conditions, it is also evident that the great majority of studies cited under Personality in the literature have not approached in rigorous conditions those imposed upon the subjects of this experiment. Rich (**25**), for instance, in his study of a group of University students (cited above), merely asked his subjects to control their diets during the day upon which they took the urine samples, and upon the two days preceding it. However, since each subject was requested to eat as nearly as possible the food to which he was accustomed, it is evident that all did not have the same diet. In his study of a group of 303 children, Rich (**25**) secured blood specimens several hours after a meal from each child. However, the samples were not taken a sufficient length of time after a meal to permit complete readjustment of the blood content. In Part II of the present study, in which one individual acted as subject in an extensive series of experimental sessions, the conditions necessary to truly basal determinations were rigidly observed.

Tables 1–6 give the basal determinations of inorganic phosphorus, cholesterol, sugar, chlorides, calcium, and creatinine in mg per 100 cc blood of Periods I, II and III in 19 subjects. Table 7 shows the pre- and post- determinations of inorganic phosphorus, cholesterol, sugar, and creatinine [4] in mg per 100 cc blood during a 3 hour period of relaxation in the control group of 14 subjects. Table 8 gives the personality ratings on neurotic tendency, dominance, self-sufficiency (Bernreuter Personality Inventory) and the test scores on the CAVD Intelligence Scale, Levels M–Q in 19 subjects.

In general, basal biochemical determinations of the same blood constituent performed at intervals of three weeks do not correlate significantly. However, the correlations be-

[4] The calcium and chloride determinations were omitted in the analyses of the control group and in the analyses of Part II.

TABLE 1

BASAL DETERMINATIONS OF INORGANIC PHOSPHORUS IN MG PER 100 CC BLOOD OF PERIODS I, II AND III IN 19 SUBJECTS

Subject	I	II	III
1	3.85	4.05	5.28
2	4.29	4.05	5.32
3	3.75	3.57	5.36
4	3.17	6.41	4.84
5	4.71	4.52	4.21
6	3.85	4.17	4.46
7	4.05	5.73	4.21
8	4.01	5.36	4.39
9	3.59	6.00	4.44
10	5.00	5.21	3.38
11	5.40	4.84	4.29
12	5.00	4.19	4.29
13	5.17	5.32	4.63
14	3.89	4.69	3.75
15	4.75	4.46	4.26
16	7.08	6.00	4.46
17	6.25	4.90	3.77
18	4.93	4.75	2.94
19	5.43	4.41	4.93
Ave.	4.641	4.875	4.380
S.D.	.947	.754	.611
C.V.	20.4	15.5	13.9

TABLE 2

BASAL DETERMINATIONS OF CHOLESTEROL IN MG PER 100 CC BLOOD OF PERIODS I, II AND III IN 19 SUBJECTS

Subject	I	II	III
1	165	154	140
2	188	170	165
3	181	141	141
4	165	172	165
5	206	192	163
6	196	225	168
7	152	124	129
8	171	136	157
9	227	229	206
10	205	193	179
11	150	188	127
12	190	144	122
13	136	146	138
14	175	171	165
15	171	156	132
16	175	132	157
17	177	166	170
18	182	171	176
19	153	165	147
Ave.	177.1	167.1	155.1
S.D.	22.4	28.7	21.6
C.V.	12.6	17.2	13.9

TABLE 3

BASAL DETERMINATIONS OF SUGAR IN MG PER 100 CC BLOOD OF PERIODS I, II AND III IN 19 SUBJECTS

Subject	I	II	III
1	88	100	90
2	100	109	106
3	102	90	90
4	107	94	112
5	102	113	93
6	115	99	101
7	112	95	99
8	109	95	104
9	93	95	94
10	112	101	97
11	95	100	90
12	112	120	109
13	102	94	92
14	105	100	89
15	102	100	90
16	105	90	90
17	113	112	99
18	108	112	106
19	114	107	106
Ave.	105.0	101.4	97.7
S.D.	6.6	7.9	7.9
C.V.	6.3	7.8	8.1

TABLE 4

BASAL DETERMINATIONS OF CHLORIDES * IN MG PER 100 CC BLOOD OF PERIODS I, II AND III IN 19 SUBJECTS

Subject	I	II	III
1	612	591	604
2	592	612	626
3	607	508	605
4	604	627	584
5	586	476	604
6	576	610	607
7	500	600	607
8	500	553	634
9	526	608	630
10	587	610	569
11	529	599	575
12	537	584	622
13	538	615	605
14	531	597	596
15	542	610	572
16	570	582	576
17	593	598	630
18	563	597	559
19	610	585	569
Ave.	563.3	587.6	598.6
S.D.	40.6	26.8	23.9
C.V.	7.2	4.6	4.0

* Chlorides in this table and succeeding tables calculated as sodium chloride.

TABLE 5

BASAL DETERMINATIONS OF CALCIUM IN MG PER 100 CC BLOOD OF PERIODS I, II AND III IN 19 SUBJECTS

Subject	I	II	III
1	10.40	11.81	12.24
2	9.95	11.56	11.42
3	12.15	11.28	12.20
4	11.80	14.01	12.20
5	11.98	11.75	12.20
6	10.58	9.31	10.78
7	8.86	10.75	12.44
8	10.33	11.03	11.61
9	10.82	10.54	10.29
10	13.92	10.78	9.49
11	12.37	12.20	12.54
12	12.30	12.20	12.53
13	10.90	10.78	11.03
14	11.75	11.71	11.47
15	11.35	11.12	11.52
16	11.90	12.49	12.25
17	12.44	12.25	10.62
18	9.50	11.76	11.75
19	11.03	11.22	10.78
Ave.	11.295	11.503	11.545
S.D.	1.217	.975	.768
C.V.	10.8	8.5	6.7

TABLE 6

BASAL DETERMINATIONS OF CREATININE IN MG PER 100 CC BLOOD OF PERIODS I, II AND III IN 19 SUBJECTS

Subject	I	II	III
1	1.40	1.67	1.34
2	1.36	1.72	1.34
3	1.40	1.65	2.00
4	1.69	1.80	2.03
5	1.89	1.89	1.68
6	1.89	1.53	1.45
7	1.94	1.11	1.38
8	2.10	1.12	1.73
9	1.67	1.53	1.45
10	1.97	1.61	1.53
11	1.87	1.67	1.51
12	1.77	1.84	1.50
13	1.87	1.45	1.42
14	2.11	1.77	1.54
15	2.05	1.83	1.35
16	1.44	1.61	1.39
17	1.81	1.66	1.26
18	1.76	1.66	1.50
19	1.59	1.45	1.37
Ave.	1.767	1.609	1.514
S.D.	.211	.200	.234
C.V.	11.9	12.4	15.5

TABLE 7

PRE- AND POST- DETERMINATIONS OF INORGANIC PHOSPHORUS, CHOLESTEROL, SUGAR AND CREATININE IN MG PER 10 CC BLOOD DURING A 3 HOUR PERIOD OF RELAXATION IN A CONTROL GROUP OF 14 SUBJECTS

Subject	Inorganic Phosphorus		Cholesterol		Sugar		Creatinine	
	Pre-	Post-	Pre-	Post-	Pre-	Post-	Pre-	Post-
1	4.49	3.61	143	140	95	95	1.64	1.71
2	5.17	5.14	176	177	99	105	1.62	1.67
3	4.60	5.21	225	230	98	99	1.56	1.51
4	4.29	3.22	163	159	103	101	1.91	1.92
5	5.03	5.28	180	183	90	94	1.71	1.71
6	4.55	6.20	145	156	99	94	1.23	1.00
7	4.63	4.97	174	182	99	96	1.65	1.63
8	4.36	5.81	174	157	90	95	1.25	1.00
9	3.91	6.36	188	181	99	99	1.30	1.28
10	4.97	4.29	203	205	99	97	1.57	1.45
11	4.72	4.57	143	150	100	94	1.55	1.46
12	4.29	3.37	167	145	96	94	1.50	1.50
13	4.17	2.83	196	186	116	112	1.46	1.39
14	4.17	4.08	216	205	99	103	1.39	1.34
Ave.	4.525	4.639	178.1	175.4	98.7	98.4	1.524	1.469
S.D.	.348	1.076	24.6	25.3	6.2	5.6	.184	.253
C.V.	7.7	23.2	13.8	14.4	6.2	5.7	12.1	17.2

TABLE 8

PERSONALITY TRAIT SCORES ON NEUROTIC TENDENCY, DOMINANCE, AND SELF-SUFFICIENCY ON BERNREUTER PERSONALITY INVENTORY AND SUMMATED TEST SCORES ON THE THORNDIKE CAVD INTELLIGENCE SCALE, LEVELS M TO Q, OF 19 SUBJECTS

Subject	Neurotic Tendency	Dominance	Self-Sufficiency	CAVD
1	86	5	3	150
2	79	8	16	134
3	61	41	85	84
4	68	30	65	116
5	25	87	26	94
6	27	97	99	86
7	85	15	85	133
8	70	19	36	118
9	29	73	79	72
10	5	84	92	125
11	9	98	83	78
12	57	62	62	112
13	47	55	91	145
14	7	95	98	80
15	33	86	67	55
16	47	63	86	85
17	69	6	37	166
18	11	91	84	103
19	51	60	43	131
Ave.	45.3	56.5	65.1	108.6
S.D.	26.8	33.0	28.7	32.2

tween the pre-relaxation blood and the post-relaxation blood (taken over an interval of three hours) in 14 subjects are, for cholesterol + .93 ± .03, for sugar + .79 ± .04, for creatinine + .97 ± .01 and for inorganic phosphorus + .16 ± .18. It is thus evident that the individual maintains a biochemical stability over relatively short periods of time, such as three hours, under laboratory controlled conditions, in cholesterol, sugar and creatinine. Since we cannot depend upon the biochemical stability of the individual for more than several hours at most (and then only under controlled laboratory conditions) and since personality and intelligence are assumed to be relatively stable, then it is clear that the biochemical variability of the individual precludes the use of the day to day or week to week amounts (or averages) of these blood constituents as guides to personality or to intelligence.

Table 9 shows the correlations of the basal determinations of the various blood constituents of Periods I, II and III.

TABLE 9

CORRELATIONS OF BASAL DETERMINATIONS OF PERIODS I, II AND III IN 19 SUBJECTS

	Inorganic Phosphorus	Cholesterol	Sugar	Chlorides	Calcium	Creatinine
Periods I and II...	+.07±.15	+.64±.09	+.14±.15	+.24±.15	+.20±.15	−.36±.13
Periods I and III..	−.31±.14	+.71±.08	+.56±.11	−.21±.15	−.25±.14	+.07±.15
Periods II and III.	−.25±.14	+.67±.09	+.42±.13	−.11±.15	+.49±.12	+.22±.15
Average.....	−.16±.14	+.67±.09	+.37±.13	−.08±.15	+.15±.14	−.02±.14

It is apparent that the basal determinations of the different periods do not correlate significantly (with the exception of cholesterol and perhaps also of sugar.)

Rich (**26**) (as reported above under Personality) found a correlation of + .64 between urinary phosphorus and creatinine. In an attempt to verify his findings, the basal inorganic phosphorus determinations were correlated with the basal creatinine determinations in Periods I, II and III. The coefficients of correlation were respectively − .08 ± .15, − .34 ± .14, and + .16 ± .15, averaging − .09 ± .15. There appears to be no relationship between the inorganic phosphorus content and the creatinine content in the blood of the individual.

Table 10 gives the averages, standard deviations and coefficients of variation of the three basal determinations for the nineteen subjects. The determinations were approximately three weeks apart. Since there is a correlation of about + .60 between the cholesterol determinations in the same individual from time to time, averaging the three values of each subject does not materially affect the group variability in cholesterol. In the other blood constituents, however, the procedure of averaging the three basal determinations of Periods I, II, and III reduces the standard deviations and consequently reduces the coefficients of variation. This occurs because the extreme values (due to little stability in the same individual from time to time) are cancelled.

Table 11 gives the correlations of the averages of the 3 basal determinations (see Table 10) with personality trait

TABLE 10

AVERAGES OF 3 BASAL DETERMINATIONS IN MG PER 100 CC BLOOD OF 19 SUBJECTS

Subject	Inorganic Phosphorus	Cholesterol	Sugar	Chlorides	Calcium	Creatinine
1.....	4.39	155	93	602	11.48	1.47
2.....	4.55	174	105	610	10.98	1.47
3.....	4.23	154	94	573	11.88	1.68
4.....	4.81	167	104	605	12.67	1.84
5.....	4.48	187	103	555	11.98	1.82
6.....	4.16	196	105	598	10.22	1.62
7.....	4.66	135	102	569	10.68	1.48
8.....	4.59	155	103	562	10.99	1.65
9.....	4.68	221	94	588	10.55	1.55
10.....	4.53	192	103	589	11.40	1.70
11.....	4.84	155	95	568	12.37	1.68
12.....	4.49	152	114	581	12.34	1.70
13.....	5.04	140	96	586	10.90	1.58
14.....	4.11	170	98	575	11.64	1.81
15.....	4.49	153	97	575	11.33	1.74
16.....	5.85	155	95	576	12.21	1.48
17.....	4.97	171	108	607	11.77	1.58
18.....	4.21	176	109	573	11.00	1.64
19.....	4.92	155	109	588	11.01	1.47
Ave....	4.632	166.3	101.4	583.2	11.442	1.629
S.D....	.408	21.1	6.3	13.7	.666	.126
C.V....	8.8	12.7	6.2	2.4	5.8	7.7

TABLE 11

CORRELATIONS OF THE AVERAGES OF 3 BASAL DETERMINATIONS (TABLE 25) WITH PERSONALITY RATINGS ON BERNREUTER PERSONALITY INVENTORY AND WITH SUMMATED TEST SCORES ON CAVD INTELLIGENCE TEST, LEVELS M–Q, OF 19 SUBJECTS

	Dominance	Neurotic Tendency	Self Sufficiency	CAVD
Inorganic Phosphorus.....	−.16±.15	+.32±.14	−.00±.15	+.15±.15
Cholesterol.....	+.39±.13	−.37±.13	+.09±.15	−.25±.14
Sugar.....	−.01±.15	+.24±.15	−.16±.15	+.32±.14
Chlorides.....	−.38±.13	+.84±.04	−.28±.14	+.48±.12
Calcium.....	+.04±.15	+.17±.15	−.09±.15	−.09±.15
Creatinine.....	+.48±.12	−.31±.14	+.24±.15	−.37±.13

scores and with intelligence test scores. In view of the fact that basal biochemical determinations of the same blood constituent of the same individual performed at intervals do not correlate significantly, the rather low correlations obtained in Table 11 do not mean anything. The day to day amounts (or their averages) of the blood constituents studied in this investigation are valueless as guides to personality and in-

telligence because of the variability of the individual. At any rate the averages of the 3 basal determinations do not correlate significantly with personality traits or with intelligence. The seemingly significant correlation of + .84 between neurotic tendency and chlorides is attributed to a correlation of + .73 between the basal chlorides of Period III and neurotic tendency. The correlations of the basal chlorides of Periods I and II with neurotic tendency are respectively + .29 and + .21. In view of the low correlations between basal chloride determinations (see Table 9) there would be no significance attached to any correlation of the averages of the determinations with personality traits or intelligence and an apparently high degree of relationship may turn out to be meaningless. The results of Rich (**25**) (discussed under Personality) while exceedingly interesting are rendered very doubtful when account is taken of the variability of the individual. In his study of University students, Rich secured 3 urine specimens from each subject one week apart. The values of these specimens were averaged for each subject. This statistical procedure is not valid when individual variability renders impossible the existence of correlation between biochemical determinations of the same blood constituent of the same individual performed at intervals of time. In Rich's study of the 303 delinquent children, only one blood specimen was obtained from each child.

An attempt was made to verify the results of Hammett (13) that emotionally excitable individuals tend to have high total metabolic variabilities and, conversely, that phlegmatic individuals tend to have low total metabolic variabilities. The method employed by Hammett in giving equal weight to each blood constituent in determining the total metabolic variability was employed in the present study. A criticism of this technique may lie in the fact that it is not known whether the various blood constituents contribute equally in determining total metabolic variability of the individual. The coefficients of variation for each blood constituent for each individual were calculated. The averages and standard deviations of the 3 basal determinations were calculated for

each individual for each blood constituent. From these averages and standard deviations the coefficients of variation were obtained for each individual for each blood constituent. The sum of the coefficients of variation of the various blood constituents for the same individual was taken as the total metabolic variability of the individual. The correlation between neurotic tendency and the total metabolic variability was $+ .41 \pm .13$. This tends to give support to Hammett's (**13**) view discussed above and indicates that emotional stability may be bound up with metabolic stability. Thus, while the actual daily amounts of these blood constituents may give no indication of emotional excitability, their day to day variation would appear to be more valuable in this respect.

Table 12 gives the coefficients of variation for each blood constituent for each individual for the 3 basal determinations. The sum of these coefficients is the total metabolic variability (**13**).

TABLE 12

COEFFICIENTS OF VARIATION AND TOTAL METABOLIC VARIABILITY OF THE 19 SUBJECTS

Subject	Inorg. Phosph.	Cholesterol	Sugar	Chlorides	Calcium	Creatinine	Total Metabolic Variability
1.....	14.40	6.70	5.69	1.44	6.85	9.76	44.84
2.....	12.11	6.06	3.56	2.29	6.63	11.88	42.53
3.....	19.02	12.25	6.02	8.06	3.56	14.65	63.56
4.....	27.50	1.99	7.32	2.91	7.59	7.70	55.01
5.....	4.60	9.58	7.95	1.02	1.54	5.44	30.13
6.....	5.99	11.88	6.80	2.57	6.37	11.81	45.42
7.....	16.25	9.04	7.14	8.59	13.69	23.36	78.07
8.....	12.38	9.28	5.66	9.80	4.76	24.49	66.37
9.....	21.32	4.70	1.06	7.61	2.05	5.88	42.62
10.....	18.05	5.54	6.14	2.85	16.32	11.25	60.15
11.....	9.36	16.23	4.34	5.11	1.12	8.77	44.93
12.....	8.03	18.64	4.11	5.98	1.12	8.62	46.50
13.....	5.88	3.11	4.54	5.83	.94	13.00	33.30
14.....	10.08	2.43	6.84	5.37	1.06	12.93	38.71
15.....	4.48	10.50	5.46	4.84	1.45	16.79	43.52
16.....	19.39	11.38	7.44	.85	1.98	6.37	46.41
17.....	20.39	2.68	5.93	2.70	6.94	14.70	53.34
18.....	21.35	2.54	2.25	2.98	9.66	6.54	45.32
19.....	8.46	4.83	3.31	2.87	1.64	6.20	27.31
Ave.....	13.58	7.86	5.35	4.40	5.01	11.59	47.79

Summary of Results: Part I

1. In general, with the biochemical indices used in this study, basal biochemical determinations of the same blood constituent of the same individual performed at intervals of 3 weeks do not correlate significantly. Since the biochemical stability of the individual cannot be depended upon for more than relatively short periods of time, such as several hours (and then only under laboratory controlled conditions) it is clear that this biochemical variability of the individual precludes the use of the day to day amounts (or their averages) of these blood constituents as a guide to relatively stable measures such as personality or intelligence.

2. Inorganic phosphorus content of the blood was not found to correlate significantly with creatinine content of the blood.

3. There is an indication, as shown by a correlation of $+ .41 \pm .13$, that individuals with many neurotic tendencies tend to have a high metabolic variability. This in line with the findings of Hammett (13).

Procedure: Part II

In Part II a more extensive series of repeated measurements of the blood values of one subject was performed. The conditions necessary to truly basal determinations were rigidly observed.

The subject of Part II of the experiment was the author. He led a life of even tenor throughout the experimental period of ten weeks. The routine of one day was quite like that of the next day. Food habits were very constant.

The determinations of plasma inorganic phosphorus (37), whole blood cholesterol (5), whole blood sugar (10), and whole blood creatinine (9) were employed in this part of the study.

The subject reported at the laboratory without breakfast at 9 A.M. He rested for one half hour in a Morris chair in a semi-recumbent position. A 10 cc blood sample was then taken by a technician. The subject reported at the laboratory twice weekly, usually on Wednesdays and Fridays, for a period of ten weeks making a total of twenty basal bloods obtained from him.

As in Part I, it is here also obvious that the subject was not under a laboratory controlled situation during the intervals separating the experimental sessions. While every attempt was made to preserve an unvarying routine of activities from day to day, it is, of course, impossible to state exactly what degree of success was attained in this endeavor. In order to see what the effect of a laboratory controlled situation would have upon the blood values in contrast to a non-controlled situation, on ten occasions (out of the 20 times that the subject reported at the laboratory) one and one half hour

sessions of relaxation were introduced in order to determine the effects of a passage of time during a state of rest and minimal overt muscular activity.

During these sessions the subject rested in a Morris chair in a semi-recumbent position. He spent the session either in sleep, idle revery or drowsiness. Blood, of course, was taken before and immediately after the session.

Results

Table 13 shows the averages, standard deviations and coefficients of variation of the total 20 basal determinations of inorganic phosphorus, cholesterol, sugar and creatinine in mg per 100 cc of blood in one subject. Table 14 shows pre- and post determinations during ten 1½ hour periods of relaxation in one subject.

TABLE 13

AVERAGES, STANDARD DEVIATIONS AND COEFFICIENTS OF VARIATION OF THE TOTAL 20 BASAL DETERMINATIONS OF INORGANIC PHOSPHORUS, CHOLESTEROL, SUGAR AND CREATININE IN MG PER 100 CC BLOOD IN ONE SUBJECT

Session	Inorganic Phosphorus	Cholesterol	Sugar	Creatinine
1	3.02	176	127	1.41
2	3.32	194	91	1.72
3	3.14	196	94	1.64
4	4.29	205	100	1.89
5	3.41	210	94	1.51
6	4.21	197	97	1.53
7	3.51	208	96	1.34
8	3.57	191	98	1.39
9	3.33	191	99	1.62
10	3.71	191	94	1.63
11	3.39	180	97	1.53
12	3.21	167	107	1.50
13	4.07	197	98	1.63
14	3.50	197	96	1.82
15	2.74	176	96	1.86
16	3.13	192	95	1.62
17	2.60	197	106	1.83
18	2.81	204	96	2.23
19	3.25	199	102	1.89
20	3.00	223	98	2.01
Ave.	3.361	194.6	99.1	1.680
S.D.	.441	11.8	6.8	.220
C.V.	13.1	6.1	6.9	13.1

Over relatively short periods of time, such as one and one half hours, and under laboratory controlled conditions, the individual may show biochemical stability. The correlations

TABLE 14

PRE- AND POST- DETERMINATIONS OF INORGANIC PHOSPHORUS, CHOLESTEROL, SUGAR AND CREATININE IN MG PER 100 CC BLOOD DURING TEN 1½ HOUR PERIODS OF RELAXATION IN ONE SUBJECT

Session	Inorganic Phosphorus		Cholesterol		Sugar		Creatinine	
	Pre-	Post-	Pre-	Post-	Pre-	Post-	Pre-	Post-
2	3.32	3.49	194	197	91	89	1.72	1.72
4	4.29	3.95	205	218	100	88	1.89	1.86
6	4.21	4.49	197	216	97	101	1.53	1.45
8	3.57	3.07	191	184	98	95	1.39	1.34
10	3.71	2.94	191	197	94	90	1.63	1.50
12	3.21	2.92	167	150	107	98	1.50	1.46
14	3.50	3.21	197	183	96	95	1.82	1.67
16	3.13	2.81	192	177	95	94	1.62	1.49
18	2.81	2.06	204	214	96	94	2.23	2.10
20	3.00	2.94	223	208	98	93	2.01	1.97
Ave.	3.475	3.188	196.1	194.4	97.2	93.7	1.734	1.656
S.D.	.464	.632	13.4	20.3	4.0	3.8	.243	.239
C.V.	13.4	19.8	6.8	10.4	4.1	4.1	14.0	14.4

between the ten pre-relaxation bloods and the ten post-relaxation bloods (taken over an interval of one and one-half hours) in one subject are for cholesterol $+ .79 \pm .08$, for sugar $+ .42 \pm .18$, for creatinine $+ .98 \pm .01$ and for inorganic phosphorus $+ .86 \pm .05$.

It is evident from Table 13 that cholesterol and sugar show the least variability as revealed by the small coefficients of variation. This tends to confirm the view arrived at in Part I that these constituents show fair stability in the individual.

A comparison of the coefficients of variation of the total basal determinations of inorganic phosphorus, cholesterol, sugar and creatinine in one subject (Table 13) with the averages of the coefficients of variation in 19 subjects (Table 12) in these blood constituents reveals the fact that the one subject of Part II was about as variable as the average subject of the experimental group of 19 subjects in Part L. It was thus assured that the one subject showed no undue metabolic instability.

The variability of this individual subject was then compared with the differences in the basal determinations of the

19 individuals who acted as the experimental group in Part I. The attempt was made to see whether the differences among the 19 subjects at any one time were greater than the variability of the single subject at different times. The averages, standard deviations and coefficients of variation of the 19 basal determinations of the experimental group of Periods I, II and III were computed for each blood constituent. Table 15 gives this data and also the data for the single sub-

TABLE 15

COMPARISON OF AVERAGES, STANDARD DEVIATIONS AND COEFFICIENTS OF VARIATION * OF 20 BASAL DETERMINATIONS OF INORGANIC PHOSPHORUS, CHOLESTEROL, SUGAR AND CREATININE IN MG PER 100 CC BLOOD IN ONE SUBJECT WITH THOSE OF THE BASAL DETERMINATIONS OF 19 SUBJECTS OF PERIODS I, II AND III

	Inorganic Phosphorus			Cholesterol			Sugar			Creatinine		
	Ave.	S.D.	C.V.	Ave.	S.D.	C.V.	Ave.	S.D.	C.V.	Ave.	S.D.	C.V.
One Subject.....	3.361	.441	13.1	194.6	11.8	6.1	99.1	6.8	6.9	1.680	.220	13.1
19 Subjects Period I.....	4.641	.947	20.4	177.1	22.4	12.6	105.0	6.6	6.3	1.767	.211	11.9
19 Subjects..... Period II.....	4.875	.754	15.5	176.1	28.7	17.2	101.4	7.9	7.8	1.609	.200	12.4
19 Subjects..... Period III.....	4.380	.611	13.9	155.1	21.6	13.9	97.7	7.9	8.1	1.514	.234	15.5

* The variation of the single basal determinations among the 19 subjects for Periods I, II and III separately is used as the measure of group variability.

ject for comparison. The variation of the single basal determinations among the 19 subjects for Periods I, II and III separately is used as the measure of group variability for comparison with the individual's variability.

It is evident from Table 15 that inorganic phosphorus,[5] sugar,[6] and creatinine coefficients of variation of the single

[5] Peters and Van Slyke (22, p. 1109) state that "even during complete rest and starvation urine phosphate output varies during the 24 hour period. . . . These excretory fluctuations parallel, in direction, changes of blood phosphate."

[6] Peters and Van Slyke (22, p. 107) in referring to the great variation in normal blood sugar state that "when one considers the numberless extraneous factors which influence the normal blood sugar during the fasting state, it is not surprising to find such a high degree of variation. Most investigators have noted, furthermore, that the variation in a given individual is almost as great as that encountered in any group of normals."

individual do not differ significantly from the coefficients of variation of the group of 19 individuals in these particular blood constituents. In these three blood constituents, one individual is about as variable as a group of 19 other individuals. He no more resembles himself from day to day than he resembles others. In cholesterol,[7] however, it seems that an individual resembles himself more closely than he resembles others. He is about one half as variable in cholesterol content as the group of 19 other individuals. This confirms the finding of Part I that individuals correlate fairly highly in cholesterol values from time to time. The author can offer no explanation for this fact except to state that the ratio of the daily intake of lipoid substances to concentrations of lipoid constituents of the blood is low compared, for instance, to the ratio of the daily intake of sugar to blood sugar concentration.

The correlation between the 20 basal determinations of inorganic phosphorus and the 20 basal determinations of creatinine did not correlate significantly ($-.33 \pm .13$), again confirming the similar finding in Part I.

Summary of Results: Part II

1. In general, the biochemical variability of the individual, as determined over intervals of several days, precludes the use of the day to day amounts (or their averages) of these blood constituents as a guide to personality or to intelligence. The biochemical stability of the individual cannot be depended upon for more than relatively short periods of time, such as several hours (and then only under laboratory controlled conditions). The variability of a single individual does not differ significantly from the variability of a group of 19 individuals in the blood constituents studied in this reaserch (with the exception of cholesterol).

2. Inorganic phosphorus content of the blood was not found to correlate significantly with creatinine content of the blood.

[7] Peters and Van Slyke (22, p. 239) state that studies of all the lipoid constituents show that they are relatively more constant in a given individual than they are in a group of individuals. However, Looney and Childs (17) conclude that no single determination of cholesterol can be taken as characteristic since the variability of the individual is so great and the range so wide.

Discussion of Results

The biochemical variability of the individual from day to day would serve to explain why some investigators have found biochemical correlates for personality disorders while others have failed to corroborate these findings. Biochemical correlates that might be found would not be due to any characteristic physiological picture but rather to chance coincidental factors of individual variability.

From the results obtained in this study, the conclusion must necessarily be reached that these biochemical measures indicate not individual traits but conditions which vary in any individual from time to time. Since it cannot be established that the individual has a characteristic value, it is probable that these conditions depend upon many fac ors, each of which varies in the effects that it exerts upon the blood values from time to time. Such factors may well be physiological states of the moment, activities performed during the immediate past, diet variations, etc. It is important to note, however, that even when these conditions were held more or less constant over a period of ten weeks, the subject still exhibited great variability within the normal range of the blood constituents. In pathological conditions, the individual goes beyond the normal range in particular blood constituents. As long as the pathological condition persists, the blood picture (such as blood sugar in diabetes, calcium in tetanoid conditions, creatinine in kidney conditions, etc.) may be relatively constant and remain so until the condition is removed. In normal healthy individuals, however, there is very considerable variation within the normal range.

Conclusions

The day to day contents (or their averages) of the blood constituents studied (inorganic phosphorus, cholesterol, sugar, chlorides, calcium, creatinine) are valueless as guides to personality or to intelligence because of the variability of the individual from day to day. However, the degree of variabil-

ity of these blood constituents appears to be correlated with the personality triat of emotional stability.

(Manuscript received May 10, 1934)

References

1. Appel, K. E., Farr, C. B., and Hodes, P. J., The acid-base equilibrium of the blood in psychotic patients, with special reference to affective psychoses, *J. Nervous Mental Diseases*, 1932, **75**, 22–33.
2. Ballif, I., and Gherscovici, I., Acid-base equilibrium in nervous and mental disease, *Compt. rend. soc. biol.*, 1931, **108**, 995–6.
3. Barrett, T. M., and Serre, P., Blood analysis and sugar tolerance tests in mental disease, *J. Nervous Mental Diseases*, 1924, **59**, 561–70.
4. Bernreuter, R. G., The personality inventory, Stanford Univ., Stanford Univ. Press, 1931.
5. Bloor, W. R., The determination of cholesterol in blood, *J. Biol. Chem.*, 1916, **24**, 227–31.
6. Bloor, W. R., McQuarrie, I., Husted, C., and Patterson, H. A., The lipoids of the blood plasma in epilepsy. I. A statistical study of single determinations in 100 epileptic and 32 "normal" subjects, *J. Clin. Investigation*, 1933, **12**, 247–54.
7. Cannon, W. B., Bodily changes in pain, hunger, fear and rage, New York, Appleton, 1929, 404 pp.
8. Clark, E. P., and Collip, J. B., A study of the Tisdall method for the determination of blood serum calcium with a suggested modification, *J. Biol. Chem.*, 1925, **63**, 461.
9. Folin, O., and Wu, H., A system of blood analysis. *J. Biol. Chem.*, 1919, **38**, 81.
10. Folin, O., and Wu, H., A system of blood analysis. I. A simplified and improved method for determination of sugar, *J. Biol. Chem.*, 1920, **41**, 367.
11. Garrett, H. E., A study of the CAVD intelligence examination, *J. Ed. Research*, 1930, **21**, 103–8.
12. Gaskill, H. V., The objective measurement of emotional reactions, Genetic Psychology Monographs, 1933, **14**, 3.
13. Hammett, F. S., Observations on the relation between emotion and metabolic stability, *Amer. J. Physiol.*, 1920, **53**, 307–11.
14. Hammett, F. S., Studies of variation in the chemical composition of human blood, *J. Biol. Chem.*, 1920, **41**, 599–615.
15. Henry, G. W., and Mangam, E., Blood in personality disorders. Biochemical studies, *Arch. Neurol. Psychiatry*, 1925, **13**, 743–9.
16. Kramer, B., and Tisdall, F. F., A simple technique for the determination of calcium and magnesium in small amounts of serum, *J. Biol. Chem.*, 1921, **47**, 475–81.
17. Looney, J. M., and Childs, H. M., The blood cholesterol in schizophrenia, *Arch. Neurol. Psychiatry*, 1933, **30**, 567–79.
18. Ludlum, S. de W., Physiologic psychiatry, *M. Clin. N. America*, 1918, **2**, 895.
19. Negus, S., Analysis of the blood of idiots, *Science*, 1931, **73**, 449–50.
20. Obraszov, G. D., Alkali reserve in psycho-neurotic children. Relationship between ammonia content and free acidity of the urine of psycho-neurotic children, *Russ, J. Physiol.*, 1930, **13**, 570–5, 576–84.

21. Patrick, J. R., and Rowles, E., Intercorrelations among metabolic rate, vital capacity, blood pressure, intelligence, scholarship, personality and other measures on university women, *J. Applied Psychology*, 1933, **17**, 507–21.
22. Peters, J. P., and Van Slyke, D. D., Quantitative clinical chemistry, vol. 1, Baltimore, The Williams and Wilkins Co.; 1931, 1264 pp.
23. Pighini, G., Cholesterol and oxycholesterol in the serum in different forms of insanity, *Z. ges. Neurol. Psychiatrie*, 1911, **4**, 629–44.
24. Powers, H. D., Biochemistry in relation to intelligence, *Science*, 1931, **73**, 316.
25. Rich, G. J., A biochemical approach to the study of personality, *J. Abn. Soc. Psychology*, 1928, **23**, 158–75.
26. Rich, G. J., A relationship between phosphorus, creatinine and acidity in urinary excretion, *Proc. Soc. Exper. Biol. and Med.*, 1928, **25**, 307–9.
27. Rich, G. J., Intelligence and body chemistry, *Science*, 1931, **74**, 21.
28. Rowe, A. W., The differential diagnosis of endocrine disorders, Baltimore, The Williams and Wilkins Co., 1932, 220 pp.
29. Schube, P. G., Blood cholesterol studies in mental disease, *Amer. J. Psychiatry*, 1932, **12**, 355–68.
30. Starr, H. E., The hydrogen ion concentration of mixed saliva considered as an index of fatigue and of emotional expression and applied to the metabolic etiology of stammering, *Amer. J. Psychology*, 1922, **33**, 394.
31. Suitsu, N., The alkaline reserve of the blood of the insane, *Amer. J. Physiol.*, 1920, **54**, 147–52.
32. Thorndike, E. L., The measurement of intelligence, New York, Teachers College, Columbia Univ., 1926, 616 pp.
33. Van Slyke, D. D., The determination of chlorides in blood and tissues, *J. Biol. Chem.*, 1923, **58**, 523–9.
34. Weinberg, A. A., Influence of the nervous system on the excretion of creatinine, *Biochem. J.*, 1921, **15**, 306–11.
35. Weston, P. G., Analyses of blood of insane patients, *Arch. Neurol. Psychiatry*, 1920, **3**, 147–50.
36. Wilson, D. W., and Ball, E. G., A study of the estimation of chloride in blood and serum, *J. Biol. Chem.*, 1928, **79**, 221.
37. Youngburg, G. E., and Youngburg, M. V., Phosphorus metabolism. I. A system of blood phosphorus analysis, *J. Lab. Clin. Med.*, 1930, **16**, 158–66.

THE QUANTITATIVE MEASUREMENT OF LATERALITY

BY C. VAN RIPER

University of Iowa

Introduction

Most of the recent investigations of laterality have been based upon three more or less well founded assumptions. The first of these is that the peripheral asymmetry of function which we term laterality or sidedness is the result of the asymmetrical functioning of the two cerebral hemispheres.[1,2] The second assumption is that handedness as expressed in manual preference for certain activities may be shifted by environmental influence, and thus come to be at variance with the 'native' or cortical laterality.[3] This shift is accomplished by training the natively non-preferred hand in activities involving speed, strength, or accuracy. The third basic assumption is that laterality is primarily quantitative in nature and is distributed according to a bimodal curve, the right-sided mode being very much the predominant one.[4,5]

The author has shown in a previous investigation [6] that visual, kinesthetic, and script patterns when simultaneously drawn with the hands on opposite sides of a vertical board provided a test of laterality which clearly differentiated extremely right handed from extremely left handed individuals. The non-preferred hand in this situation produced mirrored patterns. An ambidextrous group of subjects was shown to

[1] Peterson, G. M.: A preliminary report on right and left handedness in the rat, *J. Comp. Psychol.*, 1931, 12, 234–250.

[2] Travis, L. E.: Speech pathology. New York: D. Appleton & Co., 1931, p. 42.

[3] Downey, J.: Laterality of function, *Psychol. Bull.*, 1933, 30, 116–118.

[4] Koch, H.: A study of the nature, measurement, and determination of hand preference, *Genet. Psychol. Monog.*, 1933, 13, 117–218.

[5] Ojemann, R. H.: Studies in handedness: I. A technique for testing unimanual handedness, *J. Educ. Psychol.*, 1930, 21, 597–611.

[6] Van Riper, C.: A new test of laterality, J. Exper. Psychol., 1934, 17, 305–313.

be divided fairly equally between right handed and left handed types of performance. Inherent in the nature of this test, as in so many other laterality tests, was the defect of non-quantitativeness. An adequate test of laterality should not only be reliable, free from training influences, and able qualitatively to distinguish right sidedness from left sidedness; it must also be able to measure degree of laterality at the same time.

Apparatus and Procedure

It was the aim of the present study to develop an apparatus and technique which would make possible the quantitative measurement of laterality. This apparatus, which we may call the critical-angle board, consisted essentially of two writing boards (W_1 and W_2) which were so hinged to the up-

right support (S) that they could be varied from a position parallel with the chest surface of the subject to a position perpendicular thereto. In this latter position the writing boards rested flat against the sides of the vertical support.

By means of two steel bars the writing boards were joined to a rider (R) which travelled in two grooves along the upper surface of the vertical support. This rider was caused to

move back and forth by fastening it to a continuous belt which passed around the crank (*C*) and beneath and around the vertical support (*S*) to be joined again to the other end of the rider. Turning the crank in one direction caused the rider to move away from the subject, thereby converging the writing boards. Turning the crank in the other direction caused the writing boards to diverge toward the position parallel with the subject's chest surface. The various angles of the boards from the parallel position were measured and expressed in terms of the position of the rider along the scale (*Sc*).

Paper from a feeder roll (*F*) was fed through guides along a wide groove in the board at the right from whence it proceeded along the back of this board to a slit immediately behind the feeder roll. It was then threaded through this slit, and finally drawn along the writing surface of W_1 to the collecting spool at the left. This arrangement was convenient in that the performance of the right hand was recorded on one side of the paper and that of the left hand on the other.

The procedure used was similar to that employed in the earlier investigation and is there described in greater detail. A kinesthetic pattern, consisting of a groove cut into a board,[7] was learned while blindfolded by tracing it with a bimanual stylus. A pencil was then placed in each hand and the subject simultaneously drew the pattern with both hands on the writing boards.

The first three drawings were made with the writing surfaces in the parallel position. The subject was instructed and encouraged to draw the pattern as quickly as possible, and the experimenter used these first three drawings to establish the optimal time limits of performance. This was done by so adjusting a metronome that the subject completed the drawing within two beats.

The usual performance of most subjects for these first three trials was the production of normal patterns with both hands. However, some individuals produced the mirrored pattern with the non-preferred hand even at this zero degree

[7] This figure is illustrated in Van Riper, *Op. cit.*, p. 308.

(the parallel) position. In order to be sure that no mere chance disorientation was involved, they were told that one hand had drawn the mirrored pattern, and cautioned to be sure to draw the normal pattern with both hands. If in spite of these instructions they still produced mirroring their critical angle of mirroring was recorded as being zero degrees.

After the metronome had been adjusted to the proper rate, the subject was instructed as follows: "You are now to draw the same pattern but this time you must be sure to finish within two beats of this metronome. I shall pronounce in rhythm to the beating of the metronome the following words: 'Ready—Go—Stop.' Moreover the boards on which you write are now converged ten degrees from their former position." The subject then drew the pattern according to the above conditions at 10°, then at 20°, and thereafter at angular intervals of ten degrees until at some angle a mirrored pattern was drawn. This angle was termed his critical angle. As a check on the accuracy of the performance a further trial was given at an angle of 10° beyond this.

After a short rest the subject repeated this entire procedure, thus providing a check on the original critical angle. In this same manner critical angles were determined for the visual pattern [8] and for the script pattern which consisted of the work 'catch' pronounced by the experimenter. A second retest on all three patterns was administered to thirty-one of the subjects after an interval of four months.

The subjects used in this investigation were thirty-five thoroughly right handed, thirty-five thoroughly left handed and thirty thoroughly ambidextrous individuals selected on the basis of their scores on the handedness questionnaire described in the previous study. In addition to these subjects, a group of thirty stutterers and a group of thirty partly right handed individuals (whose handedness indices ranged from .80 to .93) were tested on the kinesthetic and visual patterns.

[8] This figure is illustrated in Van Riper; op. cit., p. 308.

Results and Discussion

The results of this study are expressed in Tables I to III. Table I shows that the average of the critical angles (at which mirroring appeared) made by members of the thoroughly right handed group when drawing the visual pattern was 27°

TABLE I

Average Critical Angles and S.D.'S

Number of Cases	Group	Visual Pattern				Kinesthetic Pattern				Script Pattern			
		Test		Retest		Test		Retest		Test		Retest	
			S.D.		S.D.		S.D.		S.D.		S.D.		S.D.
35	R. H.	27	16	32	15	31	15	33	15	20	10	26	11
35	L. H.	27	16	30	16	35	16	35	16	23	17	25	15
30	Ambi.	69	18	75	12	70	16	76	13	65	13	67	12
30	Partly R. H.	49	25			52	26						
30	Stutt.	67	19			59	15			45	17		

for the first test (T_1) and 32° for the retest (T_2). The standard deviations of the distributions were 16° and 15° respectively. For this same group the average critical angle for the kinesthetic pattern was larger, being 31° for the first test with an S.D. of 15°, and 33° for the retest with an S.D. of 15°. The average critical angle of this group when drawing the script pattern was 20° for the first test and 26° for the retest.

The average critical angles for the left handed subjects on the visual, kinesthetic and script patterns were 27°, 35°, and 23°, respectively for the first test and 30°, 35°, and 25° for the retest. The standard deviations vary from 15° to 17°.

The average critical angles for the ambidextrous group on the visual, kinesthetic, and script patterns were 69°, 70°, and 65° respectively for the first test, and 75°, 76°, and 67° for the retest. The S.D.'s vary from 12° to 18°.

Further scrutiny of Table I discloses several important facts. (1) The right handed and left handed groups were alike in that both have small angles at which mirroring appears. They differed in that the left handed subjects produced the mirrored patterns with the right hand while the right handed subjects produced mirrored patterns with the

left. In general it may be said that the results of this study bear out the findings of the former investigation in that the non-dominant hand was the one to produce the mirrored pattern.

(2) The right and left handed individuals were distinguished from the ambidextrous individuals with regard to the size of the critical angles, the ambidextrous subjects having much larger angles of mirroring on all three patterns.

(3) The overlapping between the thoroughly right and left handed groups and the ambidextrous group is small.

(4) The group consisting of subjects having handedness indices of from .80 to .93 had an average critical angle of 49° for the visual and 52° for the kinesthetic patterns. This indicates the quantitative nature of the test since these critical angles fall between those of the pure handedness groups and those of the ambidextrous group.

(5) The group of stutterers showed an average critical angle of 67° for the visual, 70° for the kinesthetic, and 45° for the script patterns, thus showing a marked tendency to perform like the ambidextrous group. However, the average handedness index of these stutterers (.87) itself indicates that such a condition existed, hence these results can hardly be said to demonstrate an essential ambidexterity in a stutterer who was peripherally very right handed. Nevertheless several very right handed stutterers did give very high critical angles and mirrored with the right hand.

The results of this study therefore lend support to the findings of Travis [9] concerning the close relationship between confused laterality and stuttering.

(6) The patterns seem not to be of equal difficulty, the order of the patterns from least to greatest difficulty being kinesthetic, visual, and script. This held true for all three major groups and for both test and retest.

(7) The critical angle for any one pattern tended to be greater the second time the test is given. The average difference in terms of degrees for test-retest is shown in Table II.

The reliability of this critical angle at which mirroring

[9] Op. cit.

TABLE II

RELIABILITY IN PERCENTAGE OF SUBJECTS PERFORMING DIFFERENTLY ON RETEST

Difference in Degrees	S_1S_2	V_1V_2	K_1K_2	S_1S_3	V_1V_3	K_1K_3
0	38%	51%	46%	36%	26%	35%
10	44%	32%	43%	39%	51%	33%
20	14%	13%	10%	19%	23%	26%
30	3%	4%	1%	6%	0%	3%
40	0%	0%	0%	0%	0%	3%
Average Difference	8.2°	7.0°	6.6°	9.7°	9.7°	10.6°
Number of Cases	100	100	100	31	31	31

Legend: S_1, S_2, S_3 = Test, first retest, second retest on script pattern
V_1, V_2, V_3 = Test, first retest, second retest on visual pattern
K_1, K_2, K_3 = Test, first retest, second retest on kinesthetic pattern

appeared was of course of great importance. However due to limitations inherent in the nature of the test, correlation techniques could not be used. The angular units of ten degrees were not truly equal as far as difficulty was concerned. Besides the quantitative measurement in degrees, the qualitative factor as to which hand produced the mirroring was also involved in the reliability of the test. For these and other reasons it was decided to express the variability of performance in terms of percentage of subjects whose critical angle on the retest varied from that on the first test by 0, 10, 20, 30, or 40 degrees. All three major handedness groups were included in this check on reliability. Examination of Table II indicates that the critical angles for 70% of the subjects did not vary by more than 10°, and that 94% of the subjects did not vary by more than 20°. Since most of this variation, as mentioned above, was in the direction of a larger critical angle, a large share of these differences between test and retest may be attributed to learning factors.

However, as has been stressed before, the reliability must be expressed in a qualitative as well as in a quantitative manner. An individual might have produced mirroring at an angle of 30° on both first test and retest and yet have been very inconsistent in that on the first test he had produced mirroring with the right hand, and on the retest with the left. These inconsistencies along with other qualitative results are given in Table III.

TABLE III

QUALITATIVE RESULTS AND RELIABILITY

Group	Pattern	L. H. M. on both T_1 T_2	R. H. M. on both T_1 T_2	No. M. on both T_1 T_2	Inconsistancy of Performance
R. H.....	Vis.....	34	0	0	1
	Kin.....	34	0	0	1
	Scr.....	34	0	0	1
L.H.....	Vis.....	0	33	0	2
	Kin.....	0	35	0	0
	Scr.....	0	2	0	2
Ambi.....	Vis.....	9	6	5	10
	Kin.....	7	8	9	7
	Scr.....	10	8	7	8

Legend: L. H. M. = Left handed mirroring
R. H. M. = Right handed mirroring
No M. = No mirroring
T_1 = First test
T_2 = Retest

This table demonstrates that the performance typical of the right handed individual is left handed mirroring, while that typical of the left handed subject is right handed mirroring. In the ambidextrous group certain individuals produce right handed mirroring, others produce left handed mirroring, while still others show no mirroring at all and are able to draw normal patterns with both hands even at 90 degrees. Moreover the inconsistency (by which is meant a change from right to left handed mirroring) is slight except for the ambidextrous group where much inconsistency from test to retest was shown. This is perhaps to be expected since ambidexterity is of itself a sign of a lack in consistent preference.

Certain control tests were run on thirty or more right and left handed subjects to determine the effect of such factors as attention, position of the hands, type of pattern, etc. It was found that even when the subject was strongly instructed to attend to the non-preferred hand and with it to be sure to make the normal pattern, the preferred hand mirrored in only four of thirty-one subjects. The influence of extreme positions such as drawing the patterns with arms

extended to their utmost or cramped closely to the body changed the typical performance of only one subject. Pattern orientation (involving asymmetry to right or left) was shown to affect the performance of only one subject. Sandpaper patterns learned by tracing with the non-preferred hand caused mirroring with the preferred hand in two of twenty-nine cases. An interesting check on the validity of the performance was achieved by means of a test in which the subject was required to draw the mirrored image of the visual pattern presented to him. Under these conditions, the normal pattern was drawn by the non-preferred hand in all but one of thirty-one subjects.

When the experiment was originally outlined it was planned to use another series of critical angles, i.e., those at which mirroring disappeared and both hands produced normal patterns. Forty subjects were tested according to the following procedure. The boards were set at 90° (in a position perpendicular to the chest surface) and the pattern was drawn at 90°, 80°, 70° and thereafter at decreasing angular intervals of ten degrees until both hands again produced normal patterns. Retests were also run on these angles and it was found that they were not only remarkably inconsistent but seemed to be greatly dependent upon variable conditions of attention and motor set. Moreover they bore no consistent relation to the critical angles found by the other method, and did not differentiate between the various laterality groups.

Another check on the procedure showed some interesting variations due to the speed of performance. The seven subjects who mirrored at zero degrees were told to open their eyes and to watch their hands as they drew the patterns as slowly as they wished. Under these conditions none of the subjects mirrored at zero degrees. But with eyes shut and no speed required three of the seven mirrored with the non-preferred hand. When both speed and blindfold were required all seven mirrored with the non dominant hand.

A final check concerned the relation between intelligence and ability to control the bilateral production of normal patterns. Ten children from a state orphanage having I.Q.'s of

80 or below had an average angle on the visual pattern of 43°. When this is compared with the average angle (27°) of the thoroughly left handed group mainly comprised of university students, one can readily see that some other factor besides intelligence and attention was responsible for the mirroring at a certain angle.

The results of the critical angle board test indicate its quantitative nature in several ways. The extremely right and left handed groups were clearly differentiated from the ambidextrous group in a quantitative manner. A group presumably midway between them on the basis of peripheral handedness had an average critical angle which was approximately midway between the critical angles of the extreme and ambidextrous groups. Moreover, the fact that the scores made by individuals within any one group varied more or less from that average is in itself symptomatic of the quantitative nature of the test.

The test in its present form is admittedly crude and any attempt to determine a critical angle which would represent finally and forever the exact state of a person's laterality would be foredoomed to failure. The type of pattern, the method of presentation and performance, the sophistication of the subject all play some part in determining at what exact angle mirrored patterns will occur. However, the important factor in determining the critical angle is the subject's laterality and the variability of performance is low enough to make the test of service in clinical diagnosis where speech defects, mirror-writing, or reading disabilities seem to be associated with confused laterality.

The phenomena described in this investigation have interesting theoretical implications. If, as certain studies seem to indicate, the natural orientation of the two hands in simultaneous activity is isotropic with reference to the body axis,[10] then the function of the converging writing boards is to facilitate this natural orientation. If the neural organization responsible for the performance on the critical angle board

[10] Fuller, J. K.: The psychology and physiology of mirror-writing, Univ. Calif. Publ. in Psychol., 1916, 220–240.

consists of two writing 'centers,' the non-dominant one operating naturally in mirror fashion, then the usual production of normal patterns with both hands at zero degrees may be caused by the dominant center imposing its own pattern upon and through the non-dominant center's activity. According to this view, increasing the angle increases the distraction and finally removes the ability of the dominant center to impose its pattern on the non-dominant center, whereupon the natural function of the latter, mirror patterning, results.

Why degree of laterality seems roughly proportional to degree of convergence, and why stutterers often mirror simultaneously with both hands are questions which lead us even further into speculation. Until more is known of the neural organization responsible for the integration of simultaneous bimanual activity, it seems wiser merely to describe the experiment and state the results.

Summary

A laterality test was described which employed the simultaneous drawing of various types of patterns upon two writing boards which could be converged through 90° from a position parallel with the subject's chest to a position perpendicular thereto. The angle of convergence at which one hand produced mirrored patterning was shown to differentiate quantitatively between thoroughly right or left handed groups and an ambidextrous group while a group of less thoroughly right handed subjects mirrored at angles approximately midway between the angles of the former groups. A group of stutterers gave critical angles approaching those of the ambidextrous subjects. The test was shown to be fairly reliable as evidenced by average differences in critical angles from test to retest. When mirrored patterns did occur they were drawn by the non-dominant hand as indicated by handedness preference.

(Manuscript received April 23, 1934)

A COMPARISON OF FIVE METHODS OF SCORING THE GALVANIC SKIN RESPONSE

BY WILLIAM A. HUNT AND EDNA B. HUNT

Connecticut College for Women

During an investigation of the conscious correlates of the galvanic skin response, the problem of scoring the deflections arose.* This question of scoring would seem to divide itself into two aspects, which may be called somewhat arbitrarily the qualitative and quantitative approaches. In the first place, the investigator is naturally interested in whether or not a given stimulus will produce a galvanic deflection. This tendency of certain stimuli to be or not to be accompanied by deflections is an important matter for our understanding of this response. We need, first, a simple statement of its frequency of occurrence in certain situations. This is what we mean by the *qualitative* aspect. In addition to whether or not the response occurs, we desire to know how large it is. We wish a measure for its size or amount. This is the aspect we refer to as *quantitative*. One may wish to differentiate between two stimuli both of which produce deflections of the same size. In one instance the deflection occurs at each occurrence of the stimulus, in the other only 50 percent of the time. Or one may wish to distinguish between two stimuli each of which gives a deflection with the same frequency of occurrence, but one of which produces apparent deflections of greater or less magnitude than the other. The possibilities are many and make the dual scoring system necessary.

As a qualitative measure the writers used the simple frequency of occurrence of the reflex expressed as a percentage of the total number of applications of the stimulus. This gives us a simple and clear expression of the tendency of occurrence.

* The writers wish to express their gratitude to Drs. Carney Landis, C. W. Darrow, and T. W. Forbes for their assistance and suggestions.

On the quantitative side the problem is not so easy. The writers originally adopted as a quantitative statement the absolute number of ohms changed of resistance with each deflection. Thus the amount would be scored as being simply the number of ohms of deflection. The question immediately arises, however, as to whether this gives a true comparative picture or whether the absolute number of ohms should not be weighted in some manner. A clearer picture might be obtained if the amount of deflection were expressed as a percentage of the individual's basic resistance at the time the stimulus was applied. Dr. T. W. Forbes also suggested that we express the size of each deflection as a percentage of the largest deflection shown by the individual during that particular sitting, that is, of his total range of deflection. While the writers felt that these three quantitative scoring systems would agree in general, it seemed best to test them experimentally. At the same time we wished to compare the qualitative with the quantitative scoring systems, although one would expect less agreement between these last two.

Since Dr. C. W. Darrow has recently suggested that conductance offers a better measure of the response than does resistance, we have included a measure of conductance among our quantitative scoring systems (**1**). Expressing the amount of the deflection as conductance rather than resistance would not change the relative order of the scores as it involves expressing the resistance values as reciprocals and merely inverts the rank order as a whole; and, as careful inspection of Waller's data will show (**2**), the increased opportunity for clerical or arithmetical errors in this superfluous transition argues against its use. For this reason the authors have selected the *difference of conductance* as a measure, using the formula $[r/R(R - r)]$, where r equals deflection in ohms and R equals resistance level before the deflection. As Dr. Darrow has pointed out this may be simplified to the formula $1/R' - 1/R$, where R equals the original resistance level, and R' equals the resistance at the peak of the reaction.

Galvanic skin responses to 30 assorted stimulus situations were taken on 22 subjects including a group of 14 normal

individuals, 4 dementia praecox cases, and 4 cases of manic-depressive psychosis. The stimuli included association words, jokes, simple arithmetical problems, sudden sounds, colors, and completion pictures and were designed to give an adequate range of stimulus potentiality (**3**).

The set-up and apparatus has been described elsewhere (**4**). Its main advantages are high sensitivity with a practically uniform small current through all subjects regardless of the subject's resistance, and continuous quantitative records with uniform scale values for subjects having different resistance levels. The deflections obtained were scored in the five ways which we have described. The different scoring systems were then compared, using the rank-difference correlation formula for ρ. In the accompanying table there are

INTERCORRELATION OF SCORING METHODS FOR GALVANIC SKIN RESPONSE

Nor. = Normal D.P. = Dementia Praecox M.D. = Manic-Depressive		Percent of Freq. of Occurrence	Ave. No. of Ohms Per Occurrence	Ave. Percent of Basic Resistance	Ave. Percent of Total Range of Deflections
Average number of ohms per occurrence	Nor.	.86			
	D.P.	.67			
	M.D.	.88			
Average percent of basic resistance	Nor.	.86	.99		
	D.P.	.67	.99		
	M.D.	.86	.98		
Average percent of total range of deflections	Nor.	.88	.96	.97	
	D.P.	.86	.90	.88	
	M.D.	.88	.99	.99	
Average conductance difference	Nor.	.83	.97	.99	.95
	D.P.	.64	.92	.92	.79
	M.D.	.68	.82	.88	.85

presented the correlations so obtained. In this table the results for the normal, dementia praecox, and manic-depressive groups are presented separately.

Inspection of the table immediately shows that no matter which of the scoring systems is used the relative rank of the

stimuli shows little change. Let us first compare the four *quantitative* measures, absolute number of ohms per deflection, the number of ohms as a percentage of the total bodily resistance, the number of ohms as a percentage of the individual's total range of deflection, and conductance difference. For the normals, the correlation of number of ohms with percentage of bodily resistance is $+ .99 \pm .01$, with percentage of total range of deflection $+ .96 \pm .01$, with conductance difference $+ .97 \pm .01$. The coefficients for the manic-depressive group are respectively $+ .98 \pm .01$, $+ .99 \pm .01$, and $.82 \pm .04$; and for the dementia praecox group $+ .99 \pm .01$, $+ .90 \pm .04$, and $.92 \pm .01$. The other intercorrelations are comparable as the table shows. The values are so high, with the occasional possible exception of the abnormal groups, as to make the relative results of either scoring system practically identical. Our results here are even higher than those of Lauer, who has reported a correlation of $+ .899$ between the size of deflections and percentage of bodily resistance (**5**). Lauer inclines to favor the use of percentage of bodily resistance on the grounds of slightly higher reliability. His differences, however, are quite small. In view of this the present writers are not inclined to follow Lauer. Until a clearer difference in reliability is demonstrated, the method of scoring by the absolute number of ohms would seem preferable to the others. It is easier and quicker since it can be read directly from the record and does not entail further mathematical treatment. In addition it seems more understandable to the reader, as it presents an immediate and vivid measure of the phenomenon in question.[1] Since our understanding of the galvanic skin response is still somewhat tentative, it would seem best to avoid the useless and esoteric multiplication of scoring systems.

In comparing the qualitative with the quantitative method we find values that are consistently lower. The correlation between percentage frequency of occurrence and absolute

[1] In view of our tentative knowledge of the electrical phenomena involved in the galvanic skin response it is necessary to point out that these high correlations may possibly be due to the particular electrical hook-up which we have employed, and might not appear if a different circuit were used.

number of ohms is, for normals $+ .86 \pm .04$, for manic-depressives $+ .88 \pm .04$, for the dementia praecox cases $+ .67 \pm .10$. The other ρ's are comparable. We have already anticipated this difference between the qualitative and quantitative measures in our discussion when we mentioned the possibility of two stimuli producing, with the same frequency, deflections of different size, and vice versa. Inspection of the raw data showed the occasional occurrence of this phenomenon. For this reason, in spite of the general agreement between them, it would seem best to retain both the qualitative and quantitative approaches in scoring.

The criticism might possibly be raised that our high correlations are due to the unifying effect produced by our using the *average* response of the group. In order to check this, the writers have run rank order correlations between the number of ohms deflection and conductance difference for each of the individual subjects. Where our table shows a correlation for the average value of the group of + .97 for the normals, the 14 individual correlations are + .85, + .91, + .92, + .93, + .94, + .96, + .96, + .98, + .98, + .98, + .99, + .99, + .99, and + .99. The results for the two abnormal groups also show the same tendency. It is evident that our use of group averages has not seriously distorted the individual picture.

In summary, the present results indicate that the absolute number of ohms deflection involved offers a simple and adequate method of scoring the galvanic skin response. For purposes of further analysis it would seem beneficial to supplement it by some measure of frequency, such as the percentage frequency of occurrence.

(Manuscript received May 15, 1934)

References

1. Darrow, C. W., The significance of skin resistance in the light of its relation to the amount of perspiration, *J. Gen. Psychol.* (In press.)
2. Waller, A. D., Concerning emotive phenomena. Part 2. Periodic variations of conductance of the palm of the human hand, *Proc. Roy. Soc. B.* 1919, **91**, 17–31.
3. Landis, C., and Hunt, W. A., The conscious correlates of the galvanic skin response, *J. Exper. Psychol.* (In press.)
4. Landis, C., and Hunt, W. A., Word-association reaction time and the magnitude of the galvanic skin response, *Amer. J. Psychol.* (In press.)
5. Lauer, A. R., Reliability of the galvanic reflex, *Amer. J. Psychol.*, 1929, **41**, 263–270.

DISCUSSION

A FURTHER NOTE ON LEARNING WITHOUT AWARENESS OF WHAT IS BEING LEARNED

BY EDWARD L. THORNDIKE

Teachers College, Columbia University

AND

ROBERT T. ROCK, JR.

Fordham University

The experiment of Irwin, Kauffman, Prior and Weaver reported in this JOURNAL (1934, **17**, 823–827) deserves comment both because of its timeliness and because of the general importance of the use of gradual formation of a mental connection as a criterion of the absence of inference, comprehension, 'learning by ideas,' 'insight' and the like, and of sudden formation as a criterion of their presence. We congratulate them upon doing what we left undone and hope that they will investigate further the weaknesses of both these criteria. It will be instructive to apply the same critique to our other experiment. Their evidence that certain individuals to whom they gave explanations of the difference between the two types of connections nevertheless progressed only gradually toward responding correctly thereafter deserves very serious consideration. It makes clear the need of better precautions than we used or than have been customary by those using gradualness and suddenness of learning as criteria.

We cannot value as highly their conclusions that "a single unfavorable case is enough to throw upon Thorndike and Rock the entire burden of proof that their subjects were learning without knowledge of what was being learned," and that our association experiment "affords no evidence of such learning without knowledge." On the contrary, that experiment seems to us still to afford substantial evidence of such learning. So also do the results in their experiment. They would, in our opinion, have been wiser to have asserted: "The first of the Thorndike-Rock experiments does afford evidence of learning without awareness, etc. We find it hard to explain its results on the hypothesis that all the gains in per-

centages correct were due to learning with awareness or to chance. But the argument from gradualness is not so strong as it is represented to be by them."

A REJOINDER

Through the kindness of the Editor of this JOURNAL we are enabled to reply to the foregoing interesting and courteous note by Professors Thorndike and Rock. We are grateful to them for having placed our objection in its larger setting and hope to undertake the investigation they suggest as to the general reliability of the criterion of gradualness. On the other hand, with regard to the main issue of the discussion we do not find the situation at all changed. Thorndike and Rock imply that Kauffman, Prior, Weaver and the writer drew too broad a conclusion from our results, and propose an alternative to it; we believe, however, that with our data at hand we were entirely justified, and that no new arguments nor evidence are contained in their note. The whole question is: how do Thorndike and Rock know that their subjects (or our own uninstructed subjects) were so unaware of the principle allotting rewards and punishments as to have learned wholly without respect to such awareness? The one criterion of lack of awareness *during the course of learning* is admitted to have failed, and no other criterion is offered to replace it. Yet without some such control the results show merely that learning has occurred, leaving unverified the assumption of lack of awareness which is essential to their conclusion. We feel that Thorndike and Rock (and ourselves as well) do not know whether this awareness existed and that consequently, as we maintained, the burden of proof is upon those who uphold the affirmative.

FRANCIS W. IRWIN

UNIVERSITY OF PENNSYLVANIA